SPECIAL REPORT

Revised Analyses of Time-Series Studies of Air Pollution and Health

Revised Analyses of the National Morbidity, Mortality, and Air Pollution Study, Part II

Revised Analyses of Selected Time-Series Studies

HEALTH EFFECTS INSTITUTE

The Health Effects Institute, established in 1980, is an independent and unbiased source of information on the health effects of motor vehicle emissions. HEI supports research on all major pollutants, including regulated pollutants (such as carbon monoxide, ozone, nitrogen dioxide, and particulate matter) and unregulated pollutants (such as diesel engine exhaust, methanol, and aldehydes). To date, HEI has supported more than 220 projects at institutions in North America and Europe and has published over 140 research reports. Consistent with its mission to serve as an independent source of information on the health effects of motor vehicle pollutants, the Institute also engages in special review and evaluation activities.

Typically, HEI receives half its funds from the US Environmental Protection Agency and half from 28 manufacturers and marketers of motor vehicles and engines in the United States. Occasionally, funds from other public and private organizations either support special projects or provide a portion of the resources for an HEI study. Regardless of funding sources, HEI exercises complete autonomy in setting its research priorities and in reaching its conclusions. An independent Board of Directors governs HEI. The Institute's Health Research and Review Committees serve complementary scientific purposes and draw distinguished scientists as members. The results of HEI-funded research and evaluations have been used in public and private decision making.

STATEMENT

Revised Analyses of Time-Series Studies of Air Pollution and Health

Over the past decade, time-series studies conducted in many cities have contributed information about the association between daily changes in concentrations of airborne particulate matter (PM) and daily morbidity and mortality. In 2002, however, investigators at Johns Hopkins University and at Health Canada identified issues in the statistical model used in the majority of time-series studies. This HEI Special Report details attempts to address several questions raised by these discoveries. The first section addresses the impact of the issues on the HEI-funded National Morbidity, Mortality, and Air Pollution Study (NMMAPS). The second section addresses the impact on additional studies selected by the US Environmental Protection Agency (EPA). Special Panels of the Health Effects Institute contributed Commentaries on the findings.

Analyses of the health effects of air pollution must account for other time-varying factors (such as weather and unmeasured risk factors) that may affect health outcomes. Otherwise, the effects of these factors could be counted as air pollution effects. Although many methods can be used for this purpose, generalized additive models (GAMs) have been the favored method in recent years. In May 2002, NMMAPS investigators at Johns Hopkins University discovered that part of the GAM programming in the S-Plus statistical software, which they and many others had used to fit GAMs to time-series data, was not entirely appropriate for this purpose. Specifically, the default convergence criteria were not appropriate and the iterative process required to obtain effect estimates was not likely to converge. After discovering these problems, the NMMAPS investigators quickly initiated alternative analyses of their data, including use of GAMs with appropriate convergence criteria, to see how the effect estimates might change. At about the same time, investigators at Health Canada found that, under certain conditions, programming to calculate standard errors of the

regression coefficients in GAM software resulted in underestimates of the standard errors.

Concurrently, results of NMMAPS and other time-series studies were under review as part of the periodic review of the National Ambient Air Quality Standards (NAAQS) for PM. Thus understanding how these results might be changed by new analyses became a priority. As the funding sponsor for NMMAPS, HEI asked the NMMAPS investigators to prepare reports presenting their new analyses. Two NMMAPS reports were submitted to HEI: "Mortality Among Residents of 90 Cities" by Dominici and colleagues and "Morbidity and Mortality Among Elderly Residents of Cities with Daily PM Measurements" by Schwartz and colleagues. A Special Panel of the HEI Health Review Committee reviewed these reports.

In the summer of 2002, EPA identified additional key studies from the US, Canada, and Europe that were cited in the draft of the Air Quality Criteria for Particulate Matter and had used GAM in their analyses. The EPA requested that the investigators who had conducted those studies also carry out and report revised analyses. The agency asked that they (1) reanalyze the original data using the same nonparametric approach (GAMs) that was used originally, but with stricter convergence criteria; and (2) examine the sensitivity of the findings obtained with GAMs when using parametric models. The latter would also estimate more accurate standard errors.

EPA requested that HEI review the resulting short communication reports of the revised analyses and write a Commentary on the effect of different analytic approaches on the results. HEI agreed to take on this effort. A Special Panel of the HEI Health Review Committee, including members of the NMMAPS Review Panel and two additional methodologists, was formed to review the short communication reports. The Panel evaluated and interpreted changes in the original results due to the revised analyses but did not specifically evaluate the original study designs and methods.

Continued

METHODS

NMMAPS Revised Analyses

Reports of the revised NMMAPS analyses addressed both problems with the application of GAMs (settings for convergence criteria and maximum iterations, and standard error estimation) and left most other aspects of the analyses unchanged. Specifically, Dominici, Schwartz, and their colleagues carried out the following:

- *Replaced GAM functions with those using stricter convergence criteria.* These analyses were designed to correct the GAM convergence problem while acknowledging that the problem with standard error estimates was not addressed.

- *Replaced GAMs with generalized linear models (GLMs) with natural cubic splines,* using approximately the same degrees of freedom as were used in the original GAMs. These analyses were aimed at correcting problems with the standard errors and provided an alternative smoothing approach to GAMs.

Schwartz and colleagues used two additional alternatives to GAM and GLM for controlling temporal effects:

- *Penalized splines* with approximately the same degrees of freedom as in the original GAMs.

- *Case-crossover matching.* This approach was used as an alternative to GAM and GLM that might be conceptually more straightforward than the regression approaches for controlling temporal effects.

As in the original report, Dominici and colleagues applied the same model to each of the 90 cities included in the evaluation of daily mortality. The same variables and smoothing functions were used in each city to control for potential confounding, while parameter estimates and fitted smooth functions were allowed to vary from city to city. Schwartz and colleagues conducted the original and the revised analyses by fitting a city-specific model to each of the 14 cities included in the analysis of hospital admissions data and each of the 10 cities included in the evaluation of mortality. Both groups

of investigators also reevaluated the effect of including copollutants in analytic models.

As in the original report, NMMAPS investigators calculated city-specific and overall estimates of mortality effects and investigated heterogeneity among cities. They also tested the sensitivity of the results to different degrees of control for unmeasured confounders.

Other Studies

In the revised analyses conducted at the request of EPA, the investigators sought to evaluate the sensitivity of effect estimates to choice of convergence criteria and maximum iterations in GAM and to use of parametric models that allow calculation of more accurate standard errors. EPA guidelines to authors suggested fitting a parametric model to the data with approximately the same degrees of freedom as for the original nonparametric model. Because of time limitations, investigators were encouraged to submit results of additional sensitivity analyses for publication elsewhere.

RESULTS

NMMAPS Revised Analyses

Overall, for the NMMAPS data, GAMs with stricter convergence criteria and GLMs with natural cubic splines resulted in lower estimates of effect than those from the original analyses conducted with GAM and default convergence criteria.

In individual cities, the revised effect estimates for mortality typically decreased and standard errors increased. Across the 90 cities, the revised mean effect on mortality decreased substantially from 0.41% (increase per 10 $\mu g/m^3$ increase in PM_{10} concentration at lag 1) to 0.27% when using GAM with stricter criteria and to 0.21% when using GLM with natural cubic splines: an overall decrease of nearly 50%. Lags 0 and 2 had corresponding decreases. Regional patterns of effect estimates remained across the 88 cities within the contiguous United States. Because the 90 city-specific estimates usually were smaller and generally had larger standard errors with the new analyses, tests for heterogeneity of effect across the cities indicated that heterogeneity was even less likely to be present than previously.

Continued

The overall decreases in effect estimates for hospitalizations for cardiovascular diseases and for chronic obstructive pulmonary disease were smaller (approximately 8% to 10%); a small but clear association continued to be found. The effect estimate on pneumonia hospitalizations was substantially reduced. As in the original studies, revised results for PM_{10} morbidity and mortality did not change substantially when copollutants were included in the models.

Other Studies

Nineteen primary authors submitted 21 short communication reports presenting results from analyses originally reported in 37 published original articles and reports. Differences between the original and revised effect estimates varied substantially across and within studies. Overall, GAMs with stricter convergence criteria and GLMs with natural cubic splines yielded lower effect estimates but largely continued to identify an association of PM with mortality and morbidity, in particular for cardiovascular and respiratory diseases. A few investigators went beyond EPA's guidance and submitted additional sensitivity analyses. The impact of these analyses also differed across studies. No substantial impact was seen in some; in others, alternative modeling of time and weather factors resulted in substantial changes.

CONCLUSIONS

On the basis of their review, the Special Panels reached the following conclusions.

Study-Specific Conclusions

- In general, the estimates of effect in NMMAPS decreased substantially, but the qualitative conclusions did not change.

- Formal tests in NMMAPS for heterogeneity of PM effect across cities did not indicate heterogeneity. The Panel recognized, however, that the power to assess the presence of heterogeneity was low because of the generally larger city-specific standard errors. The possibility of heterogeneity therefore remains.

The overall impact of the other revised analyses included:

- While the number of studies showing an association of PM with mortality was slightly smaller, the PM association persisted in the majority of studies.

- In some of the large number of studies in which the PM association persisted, the estimates of PM effect were substantially smaller.

- In the few studies in which investigators performed further sensitivity analyses, some showed marked sensitivity of the PM effect estimate to the degree of smoothing and/or the specification of weather.

General Conclusions

- The impact of using more appropriate convergence criteria on the estimates of PM effect in the revised analyses varied greatly across the studies. In some studies, stricter convergence criteria had little impact, and in a few the impact was substantial. In no study were conclusions based on the original analyses changed in a meaningful way by the use of stricter criteria. Explanations for this variability considered by the Panel include the degree of temporal smoothing used in the original analyses, the number of smoothed terms in the models, and the degree of nonlinear collinearity (concurvity) among the smoothed terms. The relative importance of these and other explanations remains unclear.

- In general, the original PM effect estimates were more sensitive to the method used to account for temporal effects than to the convergence criteria used. Further, in the few studies in which temporal effects were extensively examined, some estimates of effect were more sensitive to the degree of smoothing of temporal effects than either the convergence criteria or the method used to account for temporal effects. In some studies the original effect estimates were largely insensitive to either the method or degree of smoothing. In several studies, however, the changes were substantial enough to result in meaningful changes in the study conclusions. In those few studies in which qualitative conclusions were

Continued

changed as a result of the different approaches to smoothing, the revised results indicated no effect of PM.

• In most studies, parametric smoothing approaches used to obtain correct standard errors of PM effect estimates produced slightly larger standard errors than did GAM. The impact of these larger standard errors on level of statistical significance of the PM effect was minor.

• Alternative approaches used to model temporal effects in the revised analyses addressed the problems of obtaining incorrect effect estimates and standard errors when using GAMs. At this time, however, no approach can be strongly preferred over another for use in this context.

• These revised analyses have renewed the awareness of the uncertainties present in estimates of short-term air pollution effects based on time-series data. Neither the appropriate degree of control for time, nor the appropriate specification of the effects of weather, has been determined for time-series analyses. In the absence of adequate biological understanding of the time course of PM and weather effects, and their interactions, the Panel recommends exploration of the sensitivity of future time-series studies to a wider range of alternative degrees of smoothing and to alternative specifications of weather variables.

Impact

Air Pollution Time-Series Studies   Compared with randomized experimental studies in which the investigator controls the intervention, findings from observational studies (such as time series) are always susceptible to uncontrolled biases and must therefore be interpreted cautiously. Observational air pollution and health studies are no exception. Uncovering inappropriate default convergence criteria in the GAM function again highlights the potential for confounding in air pollution time-series studies. As in many observational studies, avoiding confounding bias typically requires identification and specification of appropriate measures of the confounding factors as terms in a regression analysis. Determining the appropriate degree of smoothing time in air pollution time-series studies

has become a central issue. Overly aggressive smoothing may allow residual confounding, whereas inadequate smoothing may allow some or all of the air pollution effect to be incorporated into the smooth term. The best method for selecting the appropriate degree of smoothing needed to control any confounding bias remains to be determined. Furthermore, as presented in the discussion of approaches to handling time, there is no gold standard for determining the appropriate degree for smoothing. The uncertainty that these issues introduce into time-series studies has motivated ongoing work to gain much needed insight. At this time, demonstration of sensitivity, or lack of it, to a range of sensible smoothing choices seems a reasonable approach.

Statistical Software   The problem with applying GAMs has sent a cautionary note to investigators using statistical software. Clearly, the S-Plus GAM function underestimated standard errors in air pollution time-series studies, and until recently, the default convergence criteria were likely to lead to incorrect effect estimates. To their credit, investigators at Johns Hopkins continued to test their models and as a result brought the issue of default convergence criteria to light.

The nearly ubiquitous use of GAMs in time-series studies reflects one of the hazards of taking a standardized approach to analysis without verifying the detailed functioning of a given software. Clearly, as in this case, widespread use by applied biostatisticians and epidemiologists does not guarantee that a software or algorithm has no drawbacks. Looking ahead, analysts need to ensure that statistical software is appropriate for a given application. Again, the use of sensitivity analyses is included among these cautions (in this case addressing sensitivity, or the lack of it, in software tuning parameters and their defaults).

Impact Calculations   Common practice has come to use effect estimates from observational air pollution studies to estimate the impact of air pollution on a large population such as an entire country. If effect estimates from the NMMAPS 90 cities mortality study were applied, the revised impact would be approximately half of the estimated impact derived using the original effect estimates. This example reinforces the need to qualify estimates of

Continued

impact by specifying the assumptions and uncertainties on which the estimates are based.

Long-Term Effects Studies Some have noted that the calculated health impact of short-term air pollution based on time-series studies is substantially smaller than that of long-term air pollution based on cohort studies. Because of the vastly larger number of time-series studies performed, however, assessors of health risk from air pollution have often had more confidence in time-series results than in results from the few cohort studies. The problem with applying GAMs has involved primarily the time-series studies, however, and correction of the problem has generally decreased estimates of effect from these studies. Thus, more emphasis on cohort studies can be expected. Further, uncertainty regarding the estimates of effect obtained from time-series studies can also be expected to place additional emphasis on long-term air pollution studies, on studies of natural experiments (so-called quasi-experimental studies), and on human and animal experimental studies.

CLOSING

The Panels were impressed by the rapid reporting and comprehensive response to the discoveries by NMMAPS and other investigators regarding GAM software used in time-series studies. NMMAPS investigators conducted and reported results of additional analyses of virtually all of their previous NMMAPS research. Authors of the short communication reports were responsive to EPA's requests and completed a great deal of work in a short period of time. As with findings of the original analyses, all of the revised findings will continue to inform the regulatory process regarding PM. At the same time, these revised analyses have renewed the interest in important questions and uncertainties that should inform future time-series analyses of air pollution and health.

CONTENTS

Revised Analyses of Time-Series Studies of Air Pollution and Health

Over the past decade, time-series studies conducted in many cities have provided information about the association between daily changes in levels of particulate matter (PM*) and daily morbidity and mortality. The National Morbidity, Mortality, and Air Pollution Study (NMMAPS), funded by HEI, was designed to address concerns about bias in the selection of cities for air pollution studies. This study evaluated mortality in the 90 largest cities in the US, as well as morbidity among elderly residents of 14 cities with daily air pollution and hospitalization data. Its investigators also estimated the cumulative effect over several days of air pollution exposure on mortality among elderly residents of 10 cities with daily pollutant monitoring (Samet et al 2000a,b). In addition, the investigators estimated the overall mean effect of PM on mortality across cities and heterogeneity among cities.

Time-series analyses such as these must account for concurrent time-varying factors that may affect health outcomes so that the effects of these factors are not counted as air pollution effects. Such factors include temperature, atmospheric pressure, and humidity as well as unmeasured risk factors. Although many methods can be used to adjust for time-varying factors, generalized additive models (GAMs) have become the favored method in recent years. In May 2002, NMMAPS investigators at Johns Hopkins University discovered that part of the programming in the S-Plus statistical software, which they and many others have used to fit GAMs to the data in time-series studies, was inappropriate to analyze such data (Dominici et al 2002). Specifically, they found that the iterative process to obtain effect estimates in the S-Plus program did not converge to the true estimate of the regression coefficients. After discovering this problem, the NMMAPS investigators quickly initiated alternative analyses, including applying GAM with appropriate convergence criteria, to see how the results might change.

* A list of abbreviations and other terms appears at the end of the Investigators' Report.

At about the same time, investigators at Health Canada (Ramsay et al 2003) found that under certain conditions GAM S-Plus software programming underestimated standard errors of the regression coefficients. Because results of NMMAPS and other time-series studies were under review as part of the periodic review of the National Ambient Air Quality Standards (NAAQSs) for PM, understanding how these results might be changed by new analyses became a priority. HEI asked the NMMAPS investigators to prepare reports presenting the results of new analyses that had been undertaken to determine how effects estimates would change with different analytic methods and to obtain more accurate standard errors. A Special Panel of the HEI Review Committee, comprising members of the earlier NMMAPS Panel, reviewed these reports.

In July 2002, the US Environmental Protection Agency (EPA)'s Clean Air Scientific Advisory Committee (CASAC) met to discuss the Air Quality Criteria for PM draft document (PM criteria document), which is to be produced every five years as part of the required review of PM NAAQSs. This document is a comprehensive literature review emphasizing new findings published since the previous criteria document. Because of the recent discovery of problems with software used in most air pollution time-series studies, EPA modified plans for discussion of the draft PM criteria document. Sections of the criteria document on epidemiologic studies were not fully reviewed at the meeting. Instead, that time was devoted to discussion of the "GAM issue" and its implications for interpretation of studies of the health effects of PM exposure. NMMAPS investigators from Johns Hopkins University and Harvard University, who had submitted preliminary reports on revised analyses to HEI and EPA, presented their revised results in that session, as did several other investigators. EPA identified additional key studies from the US, Canada, and Europe cited in the criteria document that had employed GAMs, and requested that the investigators who had conducted those studies

carry out limited revised analyses. EPA asked that they (1) conduct new analyses of their original data using the same nonparametric approach but with more stringent convergence criteria; and (2) compare the sensitivity of findings obtained with nonparametric methods to findings from parametric models. Use of the parametric models was also aimed at obtaining more accurate standard errors. EPA requested that HEI review the short communication reports submitted by these investigators on their revised analyses and write a Commentary on the effects of different analytic approaches on the findings. HEI agreed to take on this effort along with its review of the revised analyses of NMMAPS II, which was already under way. A Special Panel of the HEI Review Committee, comprising members of the earlier NMMAPS Panel with added methodologists, reviewed the short communication reports.

As explained above, the NMMAPS revised reports and the short communication reports on other revised analyses arose in different ways, and consequently, they have been reviewed somewhat differently. The NMMAPS reports provide results of revised analyses of the analyses presented in the original NMMAPS II report. Because of earlier review efforts, most Panel members already had a thorough understanding of the design, conduct, and original data analyses of NMMAPS II. Their review focused on comparing new and original NMMAPS findings, what the impact of these changes might be, and what approaches might be developed to improve future analytic methods. The 21 reports on revised analyses from other time-series studies were much shorter and included revised analyses, for the most part, of a selected number of the original analyses in these studies. For this review, the Panel looked at how results differed when applying different analytic programs, rather than evaluating the original study design. Although EPA had not requested investigators to conduct additional sensitivity studies, the Panel did review how findings were affected by different modeling of potentially confounding variables and time in those studies where the investigators had reported such sensitivity analyses.

Thus this Special Report contains two sets of reports, one on revised analyses of NMMAPS II data, and one on revised analyses of other time-series studies of the health effects of particulate air pollution.

- *Revised Analyses of the National Morbidity, Mortality, and Air Pollution Study, Part II* This part of the Special Report presents results from Dominici and colleagues and Schwartz and colleagues and a Commentary on these results by a Special Panel of the HEI Health Review Committee.

- *Revised Analyses of Selected Time-Series Studies* This part of the Special Report comprises 21 short communication reports on results of revised analyses of other time-series studies, previously reported in 37 publications and selected by the EPA as contributing key information to the current draft of the PM criteria document. It also includes a Commentary by a Special Panel of the HEI Review Committee on the findings in these reports.

This Special Report also includes an HEI Statement, a short synopsis of the findings and their implications for both sets of reports.

REFERENCES

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On generalized additive models in time series studies of air pollution and health. Am J Epidemiol 156:193–203.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity on generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Samet JM, Dominici F, Zeger SL, Schwartz J, Dockery DW. 2000a. The National Morbidity, Mortality, and Air Pollution Study, Part I: Methods and Methodologic Issues. Research Report 94. Health Effects Institute, Cambridge MA.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwartz J, Zanobetti A. 2000b. The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

ABBREVIATIONS AND OTHER TERMS

EPA	Environmental Protection Agency (US)
CASAC	Clean Air Scientific Advisory Committee
GAM	generalized additive model
NAAQS	National Ambient Air Quality Standard (US)
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
PM	particulate matter

CONTRIBUTORS

Special Panel for NMMAPS

Sverre Vedal, Chair; *National Jewish Medical and Research Center*

Ben Armstrong, *London School of Hygiene and Tropical Medicine*

David Clayton, *Wellcome Trust Centre for Molecular Mechanisms in Disease*

Edo D Pellizzari, *Analytical and Chemical Sciences HEI Health Review Committee*

Nancy Reid, *University of Toronto HEI Health Review Committee*

Daniel C Tosteson, *Harvard Medical School HEI Health Review Committee*

Mary White, *US Department of Health and Human Services*

Special Panel for Selected Revised Analyses

Sverre Vedal, Chair; *National Jewish Medical and Research Center HEI Health Review Committee*

Ben Armstrong, *London School of Hygiene and Tropical Medicine*

Kiros Berhane, *University of Southern California*

David Clayton, *Wellcome Trust Centre for Molecular Mechanisms in Disease*

Edo D Pellizzari, *Analytical and Chemical Sciences HEI Health Review Committee*

Nancy Reid, *University of Toronto HEI Health Review Committee*

James M Robins, *Harvard School of Public Health*

Daniel C Tosteson, *Harvard Medical School HEI Health Review Committee*

HEI Health Review Committee

Daniel C Tosteson, Chair; *Harvard Medical School*

Ross Anderson, *St George's Hospital Medical School*

John C Bailar III, *The University of Chicago*

John R Hoidal, *University of Utah*

Thomas W Kensler, *Johns Hopkins University*

Brian Leaderer, *Yale University School of Medicine*

Thomas A Louis, *Johns Hopkins University*

Edo D Pellizzari, *Analytical and Chemical Sciences*

Nancy Reid, *University of Toronto*

William N Rom, *New York University Medical Center*

Sverre Vedal, *National Jewish Medical and Research Center*

HEI Project Staff

Cristina I Cann, *Staff Scientist*

Aaron J Cohen, *Principal Scientist*

Sally Edwards, *Director of Publications*

Jenny Lamont, *Science Editor*

Robert O'Keefe, *Vice President*

Jane Warren, *Director of Science*

Publication Assistance

Sybil Carey, *Consulting Proofreader*

Melissa Harke, *HEI Administrative Assistant*

Frederic R Howe, *Consulting Proofreader*

Carol A Moyer, *Consulting Editor*

Ruth E Shaw, *Consulting Compositor*

HEI

Revised Analyses of the National Morbidity, Mortality, and Air Pollution Study (NMMAPS), Part II

Revised Analyses of the National Morbidity, Mortality, and Air Pollution Study, Part II

Mortality Among Residents of 90 Cities

Francesca Dominici, Aidan McDermott, Michael Daniels, Scott L Zeger, and Jonathan M Samet

Departments of Biostatistics and Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland

Morbidity and Mortality Among Elderly Residents of Cities with Daily PM Measurements

Joel Schwartz, Antonella Zanobetti, and Thomas Bateson

Environmental Epidemiology Program, Harvard School of Public Health, Boston, Massachusetts

Continued

Commentary on Revised Analyses of NMMAPS II Data

A Special Panel of the Health Review Committee

Mortality Among Residents of 90 Cities

Francesca Dominici, Aidan McDermott, Michael Daniels, Scott L Zeger, and Jonathan M Samet

ABSTRACT

This report presents findings from updated analyses of data from 90 US cities assembled for the National Morbidity, Mortality and Air Pollution Study (NMMAPS*). The data were analyzed with a generalized additive model (GAM) using the *gam* function in S-Plus (with default convergence criteria previously used and with more stringent criteria) and with a generalized linear model (GLM) with natural cubic splines. With the original method, the estimated effect of PM_{10} (particulate matter 10 μm in mass median aerodynamic diameter) on total mortality from nonexternal causes was a 0.41% increase per 10 μg/m^3 increase in PM_{10}; with the more stringent criteria, the estimate was 0.27%; and with GLM, the effect was 0.21%. The effect of PM_{10} on respiratory and cardiovascular mortality combined was greater, but the pattern across models was similar. The findings of the updated analysis with regard to spatial heterogeneity across the 90 cities were unchanged from the original analyses.

INTRODUCTION

This report describes new analyses, using updated methods, of data assembled on daily air pollution and mortality for NMMAPS. Findings of this multiyear project were previously reported, both as reports of the Health Effects

Institute (Samet et al 2000b,c) and in the peer-reviewed literature (Daniels et al 2000; Dominici et al 2000a,b; Samet et al 2000a; Zeger et al 2000). After these original publications became available, researchers discovered that the results had been affected by use of the *gam* function in the S-Plus software (Insightful Corp, Seattle WA), which introduced an upward bias in effect estimate of particulate air pollution on mortality. We have described the basis of this problem (Dominici et al 2002b) and reported the findings of initial reanalyses in an interim note to the Health Effects Institute. The revised findings and the note were also presented by Drs Dominici, Zeger and Samet to the US Environmental Protection Agency at a meeting of the Clean Air Scientific Advisory Committee of the Science Advisory Board on July 18 and 19, 2002.

Bias in the estimates resulted from implementing the GAM with the default convergence criteria in the S-Plus *gam* function (version 3.4). The potential for this bias was identified through sensitivity analyses undertaken to better understand unexpected results in analyses peripheral to the NMMAPS focus. Details of how the problem was identified, and of its consequences, are provided elsewhere (Dominici et al 2002b). Further work is being conducted to explore the basis of this bias and its dependence on details of model specification. Bias from the NMMAPS application of the GAM likely reflects the difficulty of estimating the relatively weak effect of air pollution in data with several temporally correlated variables, including air pollutant levels, temperature, and humidity.

The S-Plus (version 3.4) *gam* function had been used throughout the NMMAPS project to estimate relative rates of mortality attributable to PM_{10} while controlling for time trends, weather variables, and other possible confounders. For the recalculated analyses based on GAMs, we used markedly more stringent convergence criteria (Dominici et al 2002b).

Other possible problems with using GAMs in time-series analyses of air pollution data have been recently identified. In the presence of concurvity (that is, residual nonlinear correlation in the data), the standard error of the air pollution effect is likely to be underestimated because of the approximate method used for its calculation. Ramsay and colleagues (2003) have further investigated

the implications of this standard error approximation in time-series studies of air pollution and mortality. Dominici and colleagues (2003a) have recently released a new *gam* function that calculates the asymptotically exact standard error of the air pollution effect.

We are continuing with methodologic investigations on the adequacy of GAMs for analyses of time-series data in air pollution and health and with comparisons of a GAM with fully parametric alternatives (Dominici et al 2003a). We have completed reanalyses of the data leading to the most central findings of prior NMMAPS reports and other publications. We have carried out these analyses using: (1) the *gam* function with substantially more stringent convergence criteria; and (2) a GLM with natural cubic splines, a fully parametric alternative. Additional reanalyses of the NMMAPS data are summarized elsewhere (Dominici et al 2002a, 2003b,c).

METHODS

The methods of NMMAPS have been fully described in previous reports of the Health Effects Institute (Samet et al 2000b,c). The NMMAPS project was implemented to describe the effect of particles and other air pollutants on daily mortality across the United States. The methods were intended to provide a picture of regional variation in the effect of particles and to provide a national effect estimate, if appropriate (Dominici et al 2002a). A uniform approach taken for the within-city models was based on extensive sensitivity analyses of data for Philadelphia (Kelsall et al 1997); that is, the same variables and smoothing functions were used in each city to control for possible confounding (Table 1).

To evaluate the impact of default *gam* settings on published analyses and to provide updated results, we reanalyzed the NMMAPS data with three methods: model 1, GAM with S-Plus default convergence parameters (the original analyses); model 2, GAM with greatly more stringent convergence parameters than the defaults (Dominici et al 2002b); and model 3, Poisson regression model with parametric nonlinear adjustments for confounding factors (specifically, GLM with natural cubic splines).

Model 1 corresponds to the GAM used in prior analyses (Dominici et al 2000a, 2002a; Samet et al 2000a). Model 2 is also the GAM used in previous analyses but with stricter convergence criteria and a substantially larger number of maximum iterations in the local-scoring and backfitting algorithms (Dominici and colleagues 2003b). Comparison of estimates from models 1 and 2 provides an indication of sensitivity of findings to the convergence criteria. Model 3 (GLM) is a fully parametric analog of model 2, estimated by an iteratively reweighted least squares (IRWLS) algorithm. In model 3, we replaced smoothing splines with natural cubic splines having the same degree of freedom in the smooth functions of time, temperature, dew point, and interactions between age group indicators and the smooth functions of time. In model 3 with natural cubic splines and a fixed number of degrees of freedom, the knots were equally spaced at quantiles of the distribution of each covariate. A comparison of estimates from models 2 and 3 indicated the sensitivity of findings to the statistical method selected.

We estimated the 90 city-specific relative rates of mortality from nonexternal causes associated with a 10 $\mu g/m^3$ increase in PM_{10} under models 1, 2 and 3. The 90-city specific relative rates were pooled across cities using a two-stage hierarchical model and a three-stage regional model with noninformative priors on the variance components

Table 1. Potential Confounders or Predictors in Estimation of City-Specific Relative Rates Associated with Particulate Air Pollution Levels, and Rationale for their Inclusion in the Model

Modelling of Predictors	Primary Reasons for Inclusion
Indicator variables for the three age groups	To allow for different baseline mortality rates within each age group
Indicator variables for the day of the week	To allow for different baseline mortality rates within each day of the week
Smooth functions of time with 7 *df*/yr	To adjust for long-term trends and seasonality
Smooth functions of temperature with 6 *df*	To control for the known effects of weather on mortality
Smooth functions of dew point with 3 *df*	To control for the known effects of humidity on mortality
Separate smooth functions of time (2 *df*/yr) for each age group contrast	To separately adjust for longer-term time trends within each age group

(Dominici et al 2000a, 2002a). In Appendix A, we provide an assessment of the sensitivity of findings to details of the approach used for combining estimates across cities; these analyses show that the results are robust to these choices. Markov chain Monte Carlo analyses were performed to estimate posterior distributions of all parameters of interest with the Bayesian inference using Gibbs sampling (BUGS) program (Thomas et al 1992). For comparison, city-specific estimates were also pooled using fixed effect models and random effect models with moment estimator of the between-city variance (DerSimonian and Laird 1986).

Within each city, multipollutant models provide relative rate estimates of mortality associated with exposure to each of the pollutants included in the model. These relative rates can be pooled in a univariate or multivariate fashion. In a univariate fashion, the city-specific coefficients for each pollutant are pooled separately using fixed effects, random effects, or Bayesian methods. However, univariate pooling ignores the within-city statistical correlations among relative rate estimates for the pollutants jointly included in the model. Because of this limitation, the method of multivariate pooling, which takes these correlations into account, was chosen for all multipollutant analyses. Multivariate pooling was performed by using a Bayesian two-stage multivariate normal model (Lindley and Smith 1972) implemented by the software TLNISE (Everson and Morris 2000), which allows specification of noninformative priors on the heterogeneity covariance matrix. Sensitivity analyses of the pooled effects to the specification of the prior distribution on the covariance matrix are reported in Appendix A.

RESULTS

The reanalyses presented in this report focus on the 90 cities included in the complete NMMAPS database. Results were obtained using Bayesian methods that provide posterior means and posterior intervals of the parameters of interest. These are the Bayesian analogs of point estimates and confidence intervals, respectively. The sensitivity of findings with respect to non-Bayesian alternatives has been systematically explored in all NMMAPS analyses. We report estimates of the following quantities of interest:

- national average air pollution effect (posterior mean) and its statistical uncertainty (95% posterior interval), and

- heterogeneity of the air pollution effects across the 90 cities and its statistical uncertainty. The heterogeneity

is quantified by the between-city standard deviation of true city-specific air pollution effects.

NATIONAL AVERAGES

The national average estimates varied among models 1, 2, and 3. When we imposed stricter convergence criteria on the *gam* function of S-Plus, the national average estimate across the 90 cities at lag 1 changed from a 0.41% increase in total mortality per 10 µg/m^3 increase in PM_{10} (under model 1) to a 0.27% increase (under model 2). When GLM with natural cubic splines was used (model 3), the national average estimate was 0.21% per 10 µg/m^3 increase in PM_{10}.

Figure 1 plots the original calculations (obtained under model 1) against the updated city-specific effect estimates (obtained under model 3) for total mortality as the outcome and PM_{10} at lag 1. Note the upward bias in the original estimates and the generally close correlation between the pairs of effect estimates from model 1 and model 3. The black square is plotted at the original versus the updated national average estimate across the 90 cities. Its width and height correspond to two standard errors of the national average estimates. With the original approach, the national average estimate was 0.41% per 10 µg/m^3 increase in PM_{10} (posterior SE equal to 0.06); the updated national average estimate was 0.21% per 10 µg/m^3 increase in PM_{10} (posterior SE equal to 0.06).
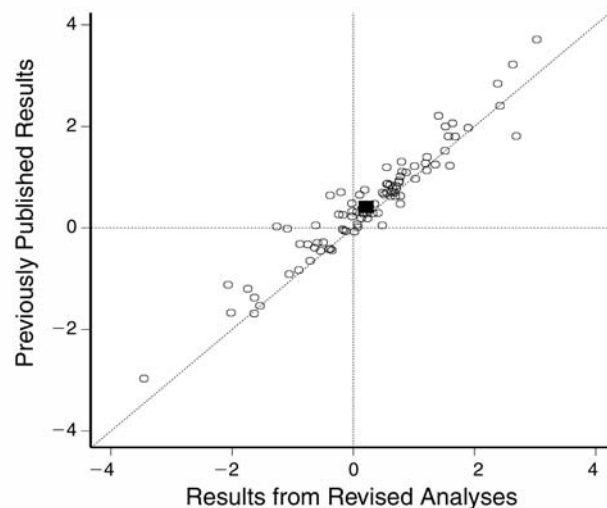
Figure 1. Percentage of change in total mortality from nonexternal causes per 10 µg/m^3 incease in PM_{10} at lag 1: previously published versus revised estimates, 90 US cities (1987–1994). The black square is plotted at the original versus updated national average estimate across the 90 cities. Its width and height corrrespond to two standard errors of the national average estimates.

Figure 2. Marginal posterior distributions for the national average effect of PM$_{10}$ on total mortality from nonexternal causes at lags 0, 1 and 2 for the 90 US cities. *A.* Revised national average estimates. *B.* Previously published national average estimates. City-specific estimates were pooled under a two-stage normal–normal hierarchical model with noninformative prior on the heterogeneity variance. Boxes provide the posterior probabilities that the national average estimates are greater than 0.

Figure 3. Posterior means and 95% posterior intervals of the national average estimates of PM$_{10}$ effects on mortality from nonexternal causes for the previously published (old) and the revised (new) estimates at lag 0, 1 and 2 for the 90 US cities.

Figure 4. Marginal posterior distributions for revised national average effects of PM$_{10}$ at lag 1 for total mortality, cardiovascular-respiratory mortality, and other causes mortality for the 90 US cities. The box at the top right provides the posterior probabilities that the overall effects are greater than 0. City-specific estimates were pooled under a two-stage normal–normal hierarchical model with noninformative prior on the heterogeneity variance.

Figure 2 provides the marginal posterior distributions of the national average estimates, original and updated, for total mortality at lags 0, 1, and 2. The national average estimates for the change in total mortality at lags 0, 1, and 2 were 0.07% (posterior SE equal to 0.06), 0.21% (posterior SE equal to 0.06), and 0.10% (posterior SE equal to 0.06). The upward bias in the original national average estimates is evident. In the updated analyses, the effect remained greatest at lag 1, and the posterior probabilities for a national average effect greater than zero were all close to one. The corresponding point estimates and 95% posterior intervals are given in Figure 3.

Figure 4 shows that the general pattern for cause-specific mortality is unchanged with the greatest effect for cardiovascular and respiratory mortality (0.31% per 10 µg/m^3 increase in PM$_{10}$ at lag 1, posterior SE equal to 0.09). The national average estimates for changes in cardiovascular and respiratory mortality at lags 0 and 2 were 0.13% (posterior SE equal to 0.09) and 0.20% (posterior SE equal to 0.09). The general insensitivity of the

Figure 5. Marginal posterior distributions for the revised national average estimates of PM_{10} on total mortality from nonexternal causes at lag 1 with and without control for other pollutants for the 90 US cities. The box at the top right provides the posterior probabilities that the overall effects are greater than 0. City-specific estimates were pooled in a multivariate fashion under a two-stage normal–normal hierarchical model.

PM_{10} effect on total mortality at lag 1 to inclusion of other pollutants in the model is shown in Figure 5. The posterior mean of the national average effect of PM_{10} on total mortality was essentially unchanged with the inclusion of either ozone (O_3) alone or of O_3 with additional pollutants. We also performed sensitivity analyses of the multivariate pooling methods with respect to prior distributions on the between-city covariance matrix. Results are summarized in Table A.1. The national average effect of PM_{10} adjusted for other pollutants was also robust to the choice of priors when pooling was performed in a multivariate fashion.

REGIONAL AVERAGES

The average estimates for total mortality are mapped in Figure 6, and the individual city estimates are plotted by region in Figure 7. The general patterns were unchanged, and the northeast continued to have the highest regional mean. Sensitivity of regional average estimates to lag specification is presented in Figure 8. Values tend to be highest within most regions at lag 1, as found with the national average estimate (Figure 2). Similar analyses were also carried out for cardiovascular and respiratory mortality (Figure 9). For this cause-of-death grouping, the pattern of regional variation was comparable to that for all nonexternal causes of death (Figure 6). The effect of PM_{10} was greatest in the northeast region. Posterior means and 95%
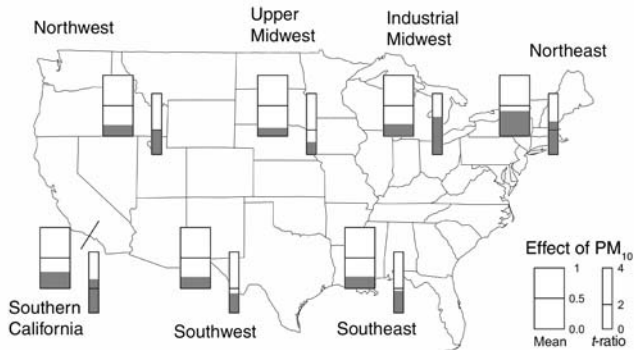
Figure 6. Posterior means divided by posterior standard deviations (*t* ratios) of regional effects of PM_{10} at lag 1 for total mortality from non-external causes.

Table 2. Posterior Means and Posterior 95% Intervals of the Between-City Standard Deviations for Total Mortality at Lag 1 Under the Three Models

	Posterior Mean	Posterior 95% Interval
Model 1[a]	0.112	0.022, 0.298
Model 2[b]	0.088	0.021, 0.240
Model 3[c]	0.075	0.021, 0.198

[a] GAM with default convergence parameters.

[b] GAM with more stringent convergence parameters.

[c] GLM with natural cubic splines.

posterior intervals of the regional average effects for total mortality and for cardiovascular and respiratory mortality are summarized in Table A.2.

HETEROGENEITY

We first tested the hypothesis of no heterogeneity by performing a standard χ^2 test (Hedges and Olkin 1985), and we accepted the null hypothesis. This result is likely due to the large statistical uncertainty of the city-specific relative risk estimates. We then repeated the test by gradually reducing the statistical variances and found that a reduction of at least 30% in the statistical variances would be necessary to reject the null hypothesis of no heterogeneity. We then implemented a Bayesian analysis with a two-stage hierarchical model and estimated marginal posterior distribution of the heterogeneity parameter. Table 2 summarizes posterior means and posterior intervals of the between-city standard deviations for total mortality and PM_{10} at lag 1 under models 1, 2 and 3. With the updated method, we estimated slightly less heterogeneity of the air
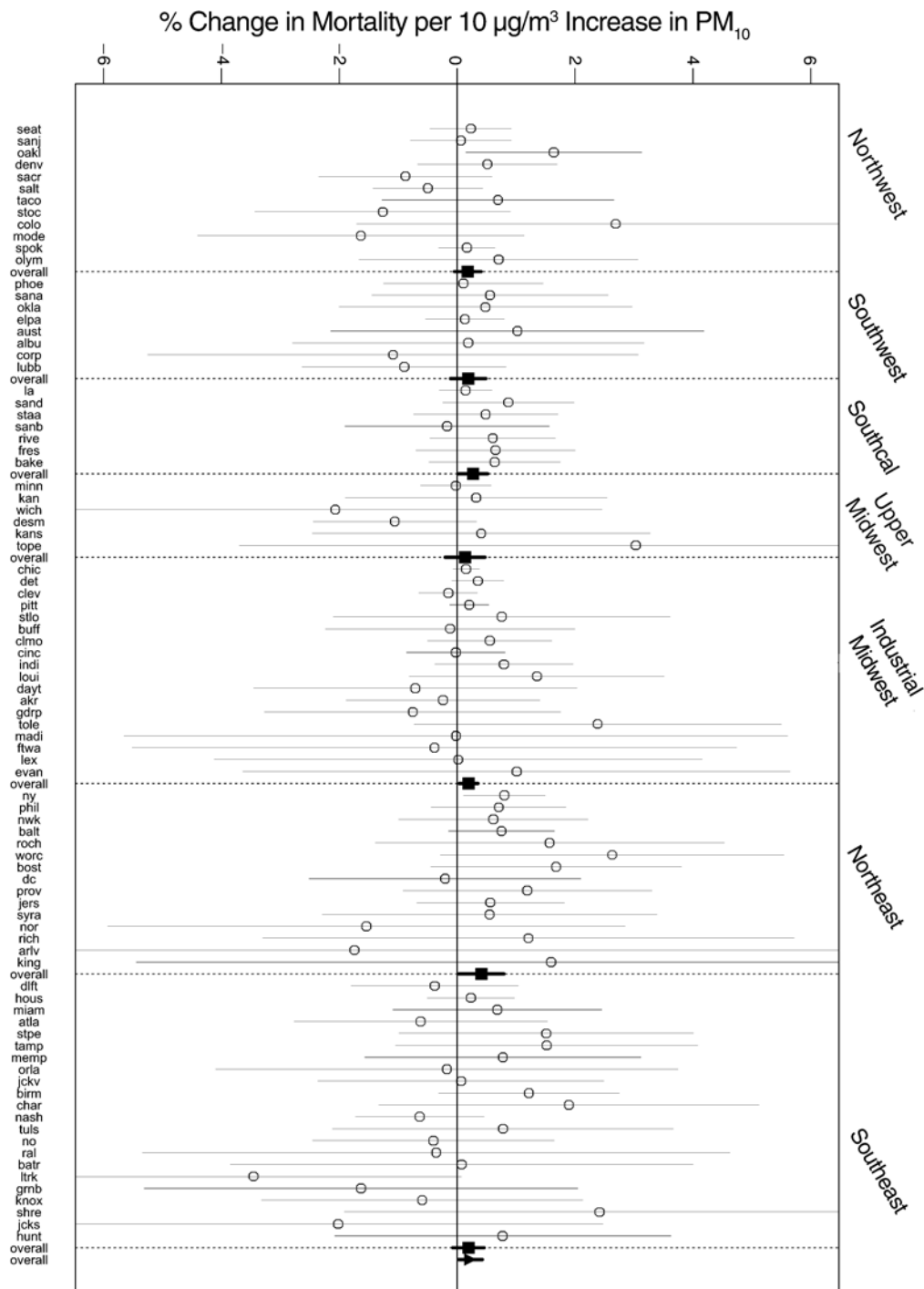
Figure 7. Maximum likelihood estimates and 95% confidence intervals of the percentage increase in total mortality from nonexternal causes per 10 μg/m³ increase in PM₁₀ for each location. The solid squares with the bold segments denote the posterior means and 95% posterior intervals of the pooled regional effects. At the bottom, marked with a triangle, is the overall effect for PM₁₀ for 88 US cities (Honolulu and Anchorage are excluded).
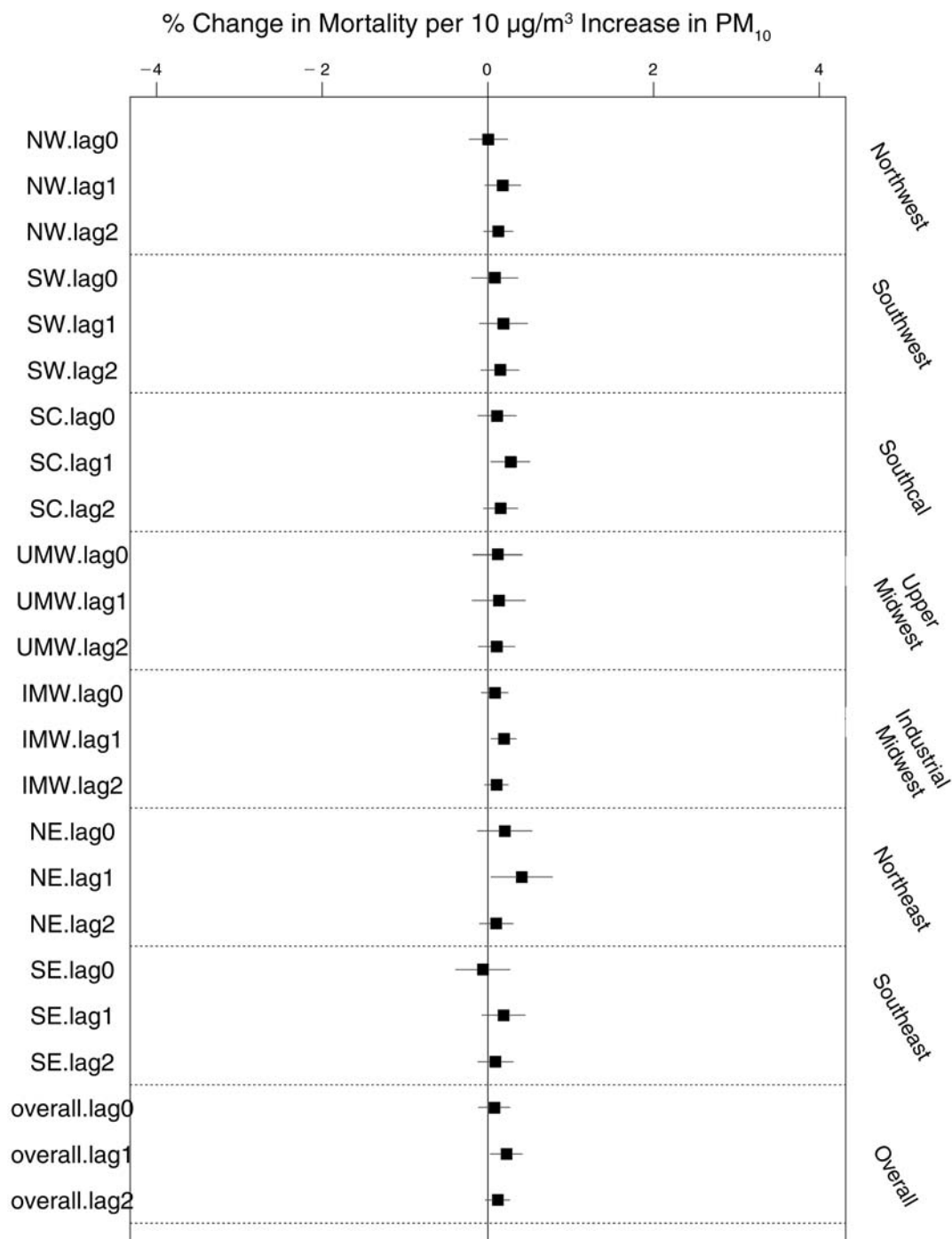
Figure 8. Posterior means and 95% posterior intervals of regional effects of PM$_{10}$ on total mortality from nonexternal causes at lags 0, 1, and 2 for the 88 US cities. At the bottom are the overall effects at lags 0, 1, and 2. City-specific estimates were pooled under a three-stage normal–normal hierarchical model with no informative priors on the within-region and between-region heterogeneity variances (see Appendix for details and sensitivity analyses to the prior distributions).
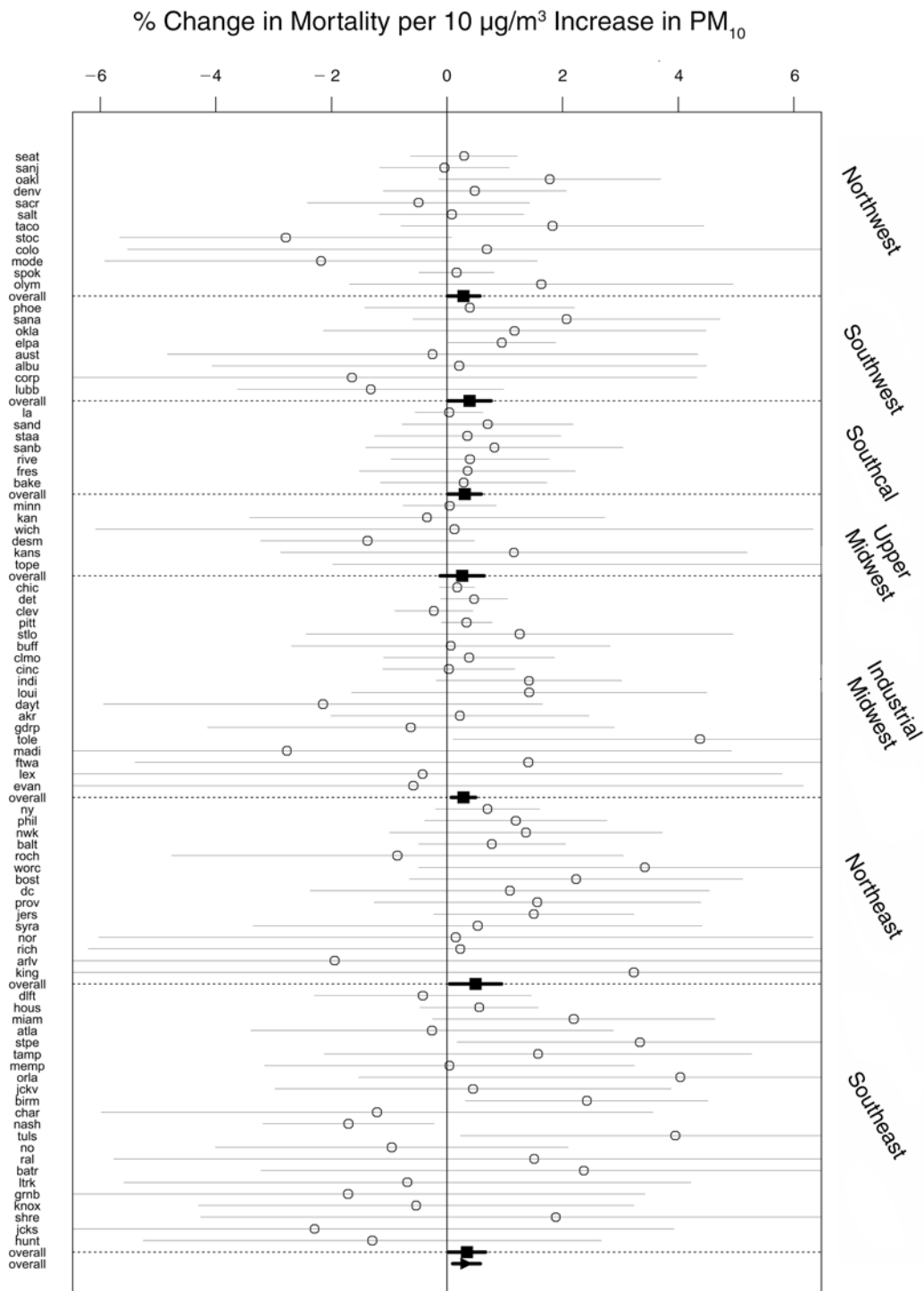
Figure 9. Maximum likelihood estimates and 95% confidence intervals of the percentage change in cardiovascular and respiratory mortality per 10 μg/m³ increase in PM₁₀ for each location. The solid squares with the bold segments denote the posterior means and 95% posterior intervals of the pooled regional effects. At the bottom, marked with a triangle, is the overall effect for PM₁₀ for the 88 US cities (Honolulu and Anchorage are excluded).
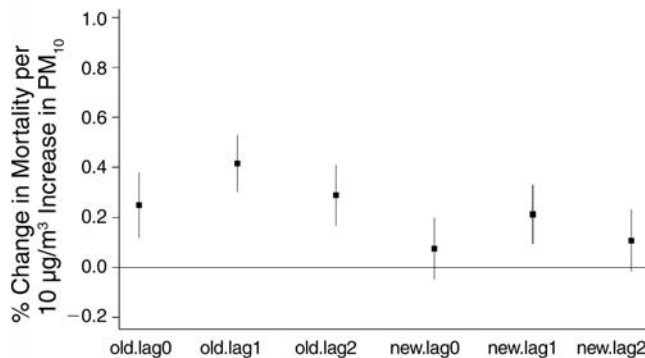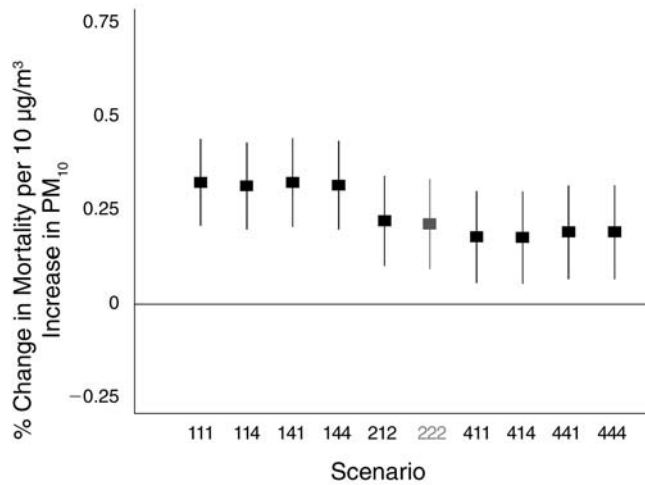
Figure 10. Posterior means and 95% posterior intervals of national average estimates of PM_{10} effects on total mortality from nonexternal causes at lag 1 for the 90 US cities, under nine alternative scenarios for adjustments for confounding factors. "222" is the baseline scenario with 7 *df* per year, 6 *df*, and 3 *df* in the smooth functions of time, temperature, and dew point, respectively. 4 and 1 denote that the *df* have been doubled and halved, respectively. Our baseline estimate (scenario "222") shows a 0.21% increase in mortality per 10 μg/m³ increase in PM_{10}. The national average estimates range from 0.32% in scenario "111" to 0.17% in scenario "414."

pollution effects among cities. Under the same model for the pooling (a two-stage normal–normal model), the posterior mean of the between-city standard deviation moved from 0.112 in model 1 to 0.088 in model 2. When GLM with natural cubic splines was used (model 3), the posterior mean of the between-city SD was 0.075. The overlap between the three posterior distributions was substantial, suggesting little sensitivity of the pattern of heterogeneity among cities to the analytic approach.

SENSITIVITY ANALYSES

We performed sensitivity analyses of the national average air pollution effects with respect to several key modeling assumptions:

- adjustment for confounding factors (Figure 10);
- choice of the prior distributions on the heterogeneity variance (eg, the between-city standard deviation squared) (Figure 11); and
- statistical models for heterogeneity: two-stage hierarchical model, three-stage regional model, and spatial correlation model (Figure A.1).

Figure 10 gives posterior means and 95% posterior intervals of the national average effects of PM_{10} for total
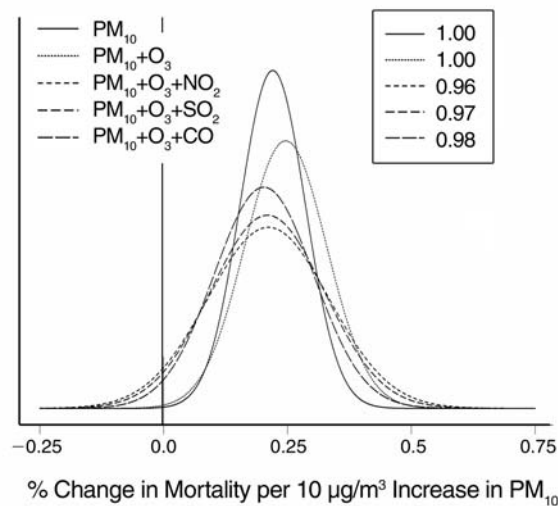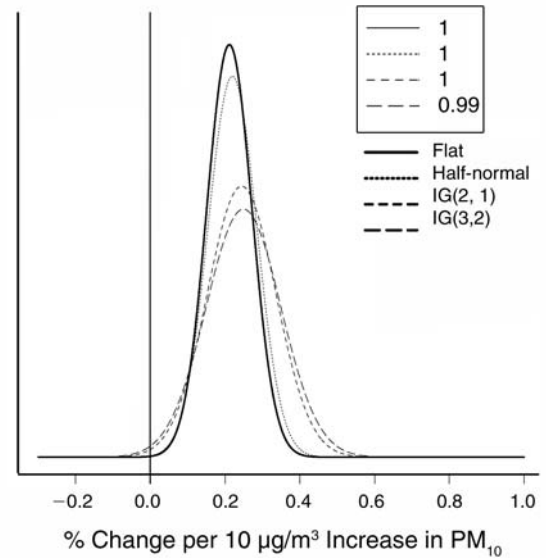

Figure 11. Marginal posterior distributions for national average estimates of PM_{10} effects on total mortality from nonexternal causes at lag 1 under four prior specifications for heterogeneity variance: 1) flat $1/\sigma^{-2}$ ~ $G(0.001, 0.001)$; 2) half-normal $\sigma^2 \sim N(0,1000)I_{\sigma^2 > 0}$; 3) $\sigma^2 \sim IG(2, 1)$; and 4) $\sigma^2 \sim IG(3, 2)$. *IG* denotes the *inverse gamma* distribution.

mortality from nonexternal causes at lag 1 under nine alternative scenarios of adjustment for confounding factors. Across the nine scenarios, we varied the degree of freedom for each of the three temporal confounders: time, temperature, and dew point. National average effects were not very sensitive to the specification of the degree of freedom in the smooth functions of time, temperature and weather. The increase in total mortality per 10 μg/m³ increase in PM_{10} of the pooled estimates ranged from 0.17% to 0.32% under the nine scenarios. The evidence of association was strong in every case.

Posterior distributions of the national average effects of PM_{10} for total mortality at lag 1 under differing prior distributions for the heterogeneity variance are plotted in Figure 11. Pooled effects of PM_{10} were only moderately sensitive to the prior distributions for the heterogeneity variance. Further sensitivity analyses of findings to the choice of the prior distributions are detailed in Appendix A. Sensitivity of the national average estimates to choice of statistical models for heterogeneity is also detailed in Appendix A.

MULTIPOLLUTANT ANALYSES

Finally, we updated analyses of the effects of pollutants other than PM_{10} on total mortality (Figures 12 through 16). These analyses include varying sets of the 90 cities, depending on the availability of the various pollutants for

the individual cities. Because the data were more limited, the 95% posterior intervals of the national averages were substantially wider for these analyses than for those directed at PM_{10}. For O_3 (Figure 12), the estimates were uniformly positive, both for O_3 alone and with inclusion of other pollutants in the model, although the 95% posterior intervals were wide. The data for estimating the effect of O_3 were limited to the summer months (Figure 13). For sulfur dioxide (SO_2) (Figure 14), nitrogen dioxide (NO_2) (Figure 15), and carbon monoxide (CO) (Figure 16), the results did not indicate associations of these pollutants with total mortality.
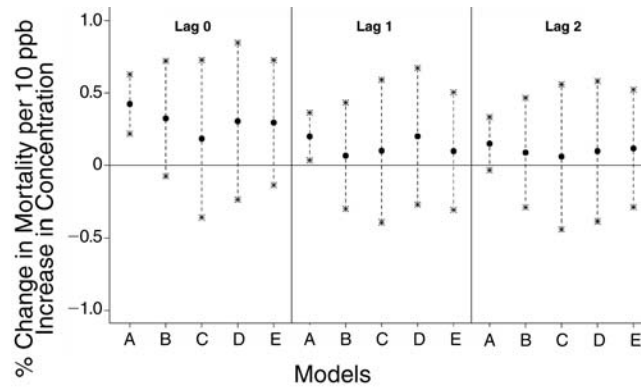
Figure 12. Posterior means and 95% posterior intervals of national average estimate of O_3 effect on total mortality from nonexternal causes at lags 0, 1, and 2 within sets of the 90 cities with pollutant data available. Models A = O_3 alone; B = O_3 + PM_{10}; C = O_3 + PM_{10} + NO_2; D = O_3 + PM_{10} + SO_2; E = O_3 + PM_{10} + CO.
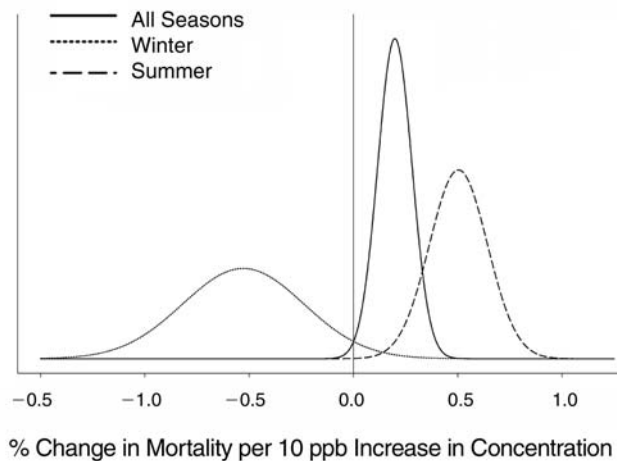
Figure 13. Marginal posterior distributions of the national average estimates of O_3 effects on total mortality at lag 0 for all seasons, summer (June, July, August) and winter (December, January, February) for the 90 cities.
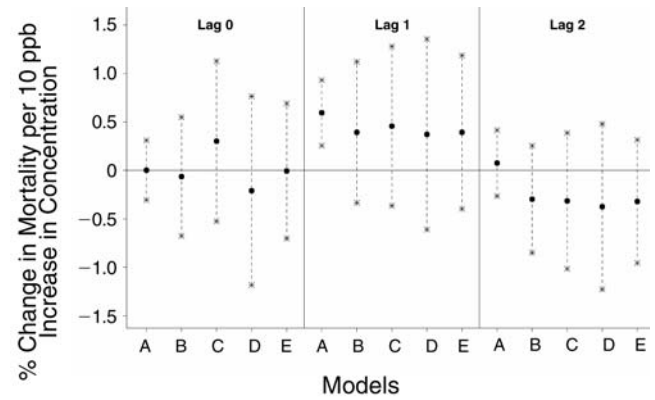
Figure 14. Posterior means and 95% posterior intervals of national average estimates of SO_2 effects on total mortality from nonexternal causes at lags 0, 1, and 2 within sets of the 90 cities with pollutant data available. Models A = SO_2 alone; B = SO_2 + PM_{10}; C = SO_2 + PM_{10} + O_3; D = SO_2 + PM_{10} + NO_2; E = SO_2 + PM_{10} + CO.
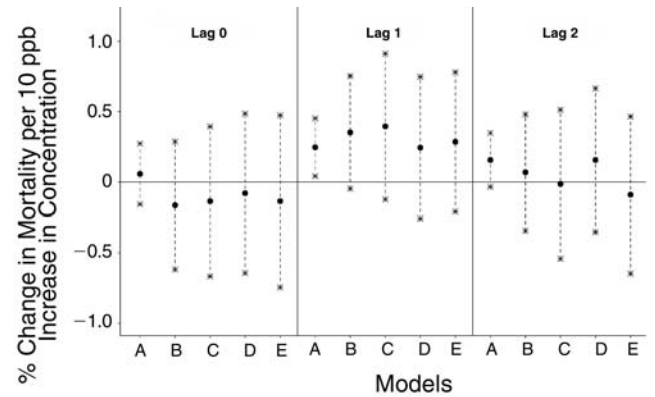
Figure 15. Posterior means and 95% posterior intervals of national average estimates for NO_2 effects on total mortality from nonexternal causes at lags 0, 1, and 2 within sets of the 90 cities with pollutant data available. Models A = NO_2 alone; B = NO_2 + PM_{10}; C = NO_2 + PM_{10} + O_3; D = NO_2 + PM_{10} + SO_2; E = SO_2 + PM_{10} + CO.
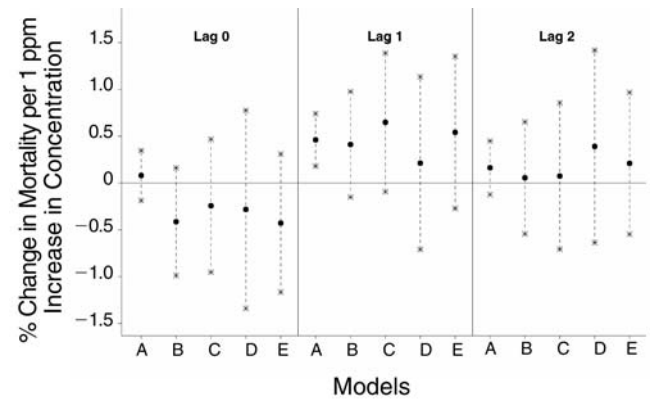
Figure 16. Posterior means and 95% posterior intervals of national average estimates for CO effects on total mortality from nonexternal causes at lags 0, 1, and 2 within sets of the 90 cities with pollutant data available. Models A = CO alone; B = CO + PM_{10}; C = CO + PM_{10} + O_3; D = CO + PM_{10} + NO_2; E = CO + PM_{10} + SO_2.

DISCUSSION

The NMMAPS modeling strategy originated with extensive exploration of the sensitivity of findings from the Philadelphia time-series data (Kelsall et al 1997). That work used GAM as the basic modeling approach and was implemented with the function *gam* in S-Plus. The *gam* default convergence criteria of S-Plus version 3.4 were used. The initial phase of model development (Kelsall et al 1997) included detailed sensitivity analyses with respect to model specification for confounder adjustment and exposure variables but not with respect to either the statistical model itself or the software for implementing the model.

We have now identified an unanticipated influence on the quantitative NMMAPS results. As documented in Dominici and colleagues (2002b), the NMMAPS estimates of national average relative risks depend upon the convergence criteria in S-Plus and on the specific statistical model used. In our particular application, reliance on the default convergence criteria led to an upward bias of 0.14% (0.41–0.27) per 10 $\mu g/m^3$ increase in PM_{10}. When a GLM was used instead of a GAM, the estimate became 0.21% per 10 $\mu g/m^3$ increase in PM_{10}. In addition, standard errors of the city-specific estimates were smaller in the GAM than in the GLM by 16% on average, even when more strict convergence parameters in the S-Plus function *gam* were used (Chambers and Hastie 1992; Ramsay et al 2003).

While the quantitative estimates changed with the tighter convergence criteria in the *gam* function or with switching to a GLM, the major scientific findings of NMMAPS did not. Strong evidence remains of an association between acute exposure to particulate air pollution (PM_{10}) and daily mortality one day later (lag 1). This association was strongest for respiratory and cardiovascular causes of death, as anticipated based on concepts of underlying susceptibility. The association of PM_{10} with mortality could not be attributed to any of the other pollutants studied: NO_2, CO, SO_2 or O_3.

In our NMMAPS reanalyses, we compared the original results to updates using a GAM with more stringent convergence criteria and a GLM. For NMMAPS, our simulation studies showed that the GLM produced less biased estimates of the pollution relative rate than did the GAM. Given our current understanding, we conclude the following.

1. GAMs with nonparametric smoothers (such as smoothing splines or locally weighted smoothers [LOESS]) provide a more flexible approach for adjusting for confounders compared to GLM with regression splines. This approach might also result in a lower prediction error for GAMs than for GLMs.

2. Work is in progress to overcome the problem of underestimating standard errors in GAM (Dominici et al 2003a); preliminary results point toward an easy and not computationally expensive solution.

3. Multicity analyses are less affected by underestimation by a GAM of the city-specific standard errors than are single-city analyses. In multicity analyses, the statistical uncertainty of the national average air pollution effect is measured by the total variance, defined as the sum of the within-city plus the between-city variance. Under hierarchical approaches for multicity analyses, underestimation of the within-city variances is compensated by overestimation of between-city variance, resulting in an almost unchanged total variance (Daniels et al 2003).

Further statistical comparisons among alternative modeling approaches for analyses of time-series data in air pollution and health are warranted, as is the development of new methods that avoid these pitfalls. In summary, the analyses reported by Dominici and colleagues (2002b) indicate some sensitivity of the quantitative, but not qualitative, results of the air pollution time-series analyses to modeling approaches and estimation procedures. Given the weak and nonspecific acute effect of air pollution on daily mortality, sensitivity of estimates to model choice would be anticipated, particularly because of the need to control for other time-correlated factors (such as temperature and season) that affect mortality. National average estimates were roughly equally sensitive to the adjustment for confounding factors under a range of plausible scenarios within model choice.

Particulate matter continued to be associated with mortality in the updated analyses, and the general pattern of spatial variability across the United States was unchanged. Ozone was associated with total mortality in the summer months. In our judgment, the new sources of uncertainty arising from model choice lead to quantitative changes in estimates without qualitative implications. Development of analytic models underlying the NMMAPS results involved multiple points of decision with assumptions: for example, choice of statistical model (eg, GAM or GLM), adjustment for confounding factors (eg, specification of temperature, season, and long-term trends in disease), and specification of exposure. Each of the decisions made by the modeler may affect the model results to a degree. While some sensitivity of findings to modeling decisions is inherent in estimating the small acute effects of particles on mortality, overall consistency of results across reasonable modeling choices is needed. We have found such consistency in these updated analyses of the NMMAPS data.

REFERENCES

Chambers JM, Hastie TJ. 1992. Statistical models in S. Chapman and Hall, London, England.

Daniels M, Dominici F, Samet JM, Zeger SL. 2000. Estimating PM_{10}-mortality dose-response curves and threshold levels: An analysis of daily time-series for the 20 largest US cities. Am J Epidemiol 152:397–406.

Daniels M, Dominici F, Zeger SL, Samet JM. 2003. Standard Error Estimation in Multi-Site Time Series Studies of Air Pollution and Health. Epidemiology (in press).

DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. Control Clin Trials 7:177–188.

Diggle P, Moyeed R, Tawn J. 1998. Model-based geostatistics. Appl Stat 47:559–573.

Dominici F, Daniels M, Zeger SL, Samet JM. 2002a. Air pollution and mortality: Estimating regional and national dose-response relationships. J Am Stat Assoc 97:100–111.

Dominici F, McDermott A, Zeger SL, Samet JM. 2002b. On generalized additive models in time series studies of air pollution and health. Am J Epidemiol 156:193–203.

Dominici F, McDermott A, Zeger SL, Samet JM. 2003a. National maps of the effects of particulate matter on mortality: Exploring geographical variation. Environ Health Perspect 111:39–43.

Dominici F, McDermott A, Zeger SL, Samet JM. 2003b. Airborne particulate matter and mortality: Time-scale effects in four US cities. Am J Epidemiol (in press).

Dominici F, Samet J, Zeger SL. 2000a. Combining evidence on air pollution and daily mortality from the largest 20 US cities: A hierarchical modeling strategy (with discussion). J R Stat Soc A 163:263–302.

Dominici F, Zeger S, Samet J. 2000b. A measurement error correction model for time-series Studies of air pollution and mortality. Biostatistics 1:157–174.

Environmental Protection Agency (US). 1996. Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information. OAQPS Staff Paper. US Government Printing Office, Research Triangle Park NC.

Everson PJ, Morris CN. 2000. Inference for multivariate normal hierarchical models. J R Stat Soc B 62:399–412.

Hedges LV, Olkin I. 1985. Statistical methods for meta-analysis. Academic Press, Orlando FL.

Kelsall JE, Samet JM, Zeger SL, Xu J. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Lindley DV, Smith AFM. 1972. Bayes estimates for the linear model (with discussion). J R Stat Soc 34:1–41.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. 2000a. Fine particulate air pollution and mortality in 20 US cities, 1987–1994. N Engl J Med 343:1742–1749.

Samet JM, Dominici F, Zeger SL, Schwartz J, Dockery DW. 2000b. The National Morbidity, Mortality, and Air Pollution Study, Part I: Methods and Methodologic Issues. Research Report 94. Health Effects Institute, Cambridge MA.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwartz J, Zanobetti A. 2000c. The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Thomas A, Spiegelhalter DJ, Gilks WR. 1992. BUGS: A program to perform Bayesian inference using Gibbs sampling. In: Bayesian Statistics 4. Proceedings of the Fourth Valencia International Meeting. Clarendon Press, Oxford, England.

Zeger S, Thomas D, Dominici F, Cohen A, Schwartz J, Dockery D, Samet J. 2000. Exposure measurement error in time-series studies of air pollution: Concepts and consequences. Environ Health Perspect 108:419–426.

APPENDIX A. Sensitivity Analyses

NATIONAL AVERAGE WITH RESPECT TO MODELS FOR HETEROGENEITY

We have performed sensitivity analyses on the estimate of national average with respect to model assumptions about heterogeneity and spatial correlation. In the first model, known as the three-stage regional model (Daniels et al 2003), we grouped the 88 counties into seven geographic regions (Northwest, Upper Midwest, Industrial Midwest, Northeast, Southern California, Southwest, Southeast), following the stratification of the United States used in the *1996 Review of the National Ambient Air Quality Standards for Particulate Matter* (Environmental Protection Agency 1996). We assumed that city-specific estimates belonging to a particular region have a distribution with mean equal to the corresponding regional effect. This assumption implies that regional heterogeneity exists: City-specific estimates of

the air pollution effects are shrunk toward their regional means, and regional means are shrunk toward the national mean, respectively. The results of this model are shown in Figures 6 and 9 and in Table A.2.

One limitation of this regional modeling approach is that two cities, far apart in terms of their geographical distance but belonging to the same geographical region, are considered more similar than two closer cities that belong to two separate geographical regions. To overcome this limitation, we relaxed the regional model assumption by developing a second model called the spatial correlation model (Dominici et al 2002a). Here we assumed that each city-specific air pollution effect is shrunk toward the average air pollution effects in the neighboring cities, where the definition of neighboring cities is based on their geographical distance (Diggle et al 1998).

The national average estimate is robust to these model assumptions. Posterior means and posterior 2.5 and 97.5 percentiles of the national average estimate for total mortality

and PM_{10} at lag 1 can be calculated under three models:

1. two-stage hierarchical model (our baseline approach applied to 90 cities), which assumes independence among the city-specific effects;

2. three-stage regional model (described previously as applied to 88 cities); and

3. spatial correlation model (described previously as applied to 88 cities).

With these models, the results were 0.21 (0.04, 0.33), 0.22 (0.03, 0.40), and 0.22 (0.10, 0.38), respectively.

Figure A.1A shows the marginal posterior distributions of the overall effects under models 1, 2 and 3. As expected, model 2 shows a slightly larger posterior interval than models 1 and 3 because of the assumption of regional heterogeneity. More specifically, in model 2 the heterogeneity is defined as total variance (ie, the sum of the variance across cities of the city-specific effects within each region plus the variance across regions of the regional estimates).

Table A.1. National Average Effects of PM_{10} at Lag 1 on Total Mortality Under Four Multipollutant Models and Four Prior Distributions on Heterogeneity Covariance Matrix[a]

	Posterior Means and 95% Posterior Regions			
	$PM_{10} + O_3$	$PM_{10} + O_3 + NO_2$	$PM_{10} + O_3 + SO_2$	$PM_{10} + O_3 + CO$
Prior 1	0.27 (0.12,0.44)	0.21 (−0.01,0.44)	0.21 (0.02,0.42)	0.24 (0.05,0.43)
Prior 2	0.25 (0.08,0.44)	0.22 (−0.04,0.50)	0.22 (−0.03,0.47)	0.20 (0.00,0.41)
Prior 3	0.28 (0.11,0.46)	0.21 (0.03,0.43)	0.20 (0.00,0.41)	0.24 (0.05,0.45)
Prior 4	0.24 (0.09,0.41)	0.21 (−0.02,0.46)	0.21 (−0.02,0.46)	0.20 (0.00,0.39)

[a] Prior 1 is a uniform prior on the shrinkage covariance matrix, Prior 2 is the reference prior, Prior 3 is the Jeffrey's prior, and Prior 4 is a flat prior on the second stage covariance matrix. See Everson and Morris (2000) for details on the prior distributions.

Table A.2. Regional Effects for PM_{10} at Lag 1 on Total Mortality from Nonexternal Causes and on Cardiovascular and Respiratory Mortality

	Posterior Means and 95% Posterior Regions		
Regions	Total (Baseline prior)	Total Mortality (Prior A)	CVDRESP (Baseline prior)
Northwest	0.18 (−0.04,0.40)	0.21 (−0.06,0.49)	0.29 (0.01,0.56)
Southwest	0.19 (−0.10,0.48)	0.22 (−0.10,0.55)	0.39 (0.02,0.75)
Southcal	0.27 (0.03,0.51)	0.30 (0.00,0.60)	0.31 (0.02,0.59)
Uppermidwest	0.13 (−0.19,0.46)	0.19 (−0.18,0.55)	0.26 (−0.11,0.63)
Industmidwest	0.19 (0.04,0.35)	0.24 (0.00,0.47)	0.29 (0.08,0.49)
Northeast	0.41 (0.04,0.78)	0.38 (0.01,0.76)	0.50 (0.05,0.94)
Southeast	0.19 (−0.07,0.45)	0.22 (−0.07,0.51)	0.34 (0.03,0.66)
National	0.22 (0.03,0.42)	0.25 (0.03,0.48)	0.34 (0.10,0.57)

Figure A.1B shows the profile likelihood (obtained under a two-stage normal–normal hierarchical model) and the marginal posterior distributions of the standard deviation among cities of the true relative rate under models 1, 2, and 3. A profile likelihood shows the weight of the evidence about the amount of heterogeneity. This profile likelihood gave the largest weights (heights of the histogram) at values close to zero, indicating homogeneity or almost no heterogeneity. The marginal posterior distributions of the between-city standard deviations under models 1 and 3 are similar to the profile likelihood. The larger posterior mean of between-city standard deviation under model 2 reflects the assumption of regional heterogeneity (that is, a larger total variance).

PRIOR DISTRIBUTIONS FOR HETEROGENEITY PARAMETERS UNDER A THREE-STAGE REGIONAL MODEL

In combining the data across cities, it is necessary to make assumptions concerning the extent of heterogeneity in the effect of air pollution on mortality among the locations. The consequences of this assumption are explored below.

Let σ^2 and τ^2 denote the variance within region and the variance across regions, respectively, of the true air pollution effects. In our baseline analyses (prior model B), we allow a range from little or no to more substantial heterogeneity, and we assume noninformative priors on σ^2 and τ^2 (gamma [0.001, 0.001] for $1/\sigma^2$ gamma [0.001, 0.001] for $1/\tau^2$).

In an alternative to our baseline analysis, which uses noninformative priors, we assume heterogeneity of the city-specific effects within regions (prior model A). More specifically, in prior model A, we allow the assumption that there is heterogeneity across cities, possibly substantial in size, and exclude within-region homogeneity (inverse gamma [3,1] for σ^2). The effect of this prior assumption with respect to prior B is twofold: It produces city-specific relative risk estimates that draw less heavily on data from each city, and it yields slightly more conservative confidence bands on the overall relative risk.

Figure A.2 shows city-specific posterior means and 95% confidence regions for each of the 88 locations under the baseline prior (prior model B—noninformative about heterogeneity) and the alternative prior distribution (prior model A—assuming heterogeneity within regions). Also shown are the posterior estimates and 95% intervals for the regional and overall means. Estimates of the overall PM_{10} relative risk are similar for the two prior models (B: 0.22 [0.02, 0.43], A: 0.25 [0.03, 0.47]). Note that under the baseline prior B, the city-specific and region-specific estimates are more like one another. This is because the posterior distributions for the within-region and between-region standard deviations of the true air pollution effects (σ and τ) are centered at mean 0.08 and 0.16 respectively, indicating a small degree of heterogeneity of the effects within a region and across regions. For example, a median value of $\sigma = 0.08$ corresponds to 95% of cities having the PM_{10} relative risks of $\pm 2 \times 0.08 = \pm 0.16$ or approximately $\pm 40\%$ of the overall relative risk.
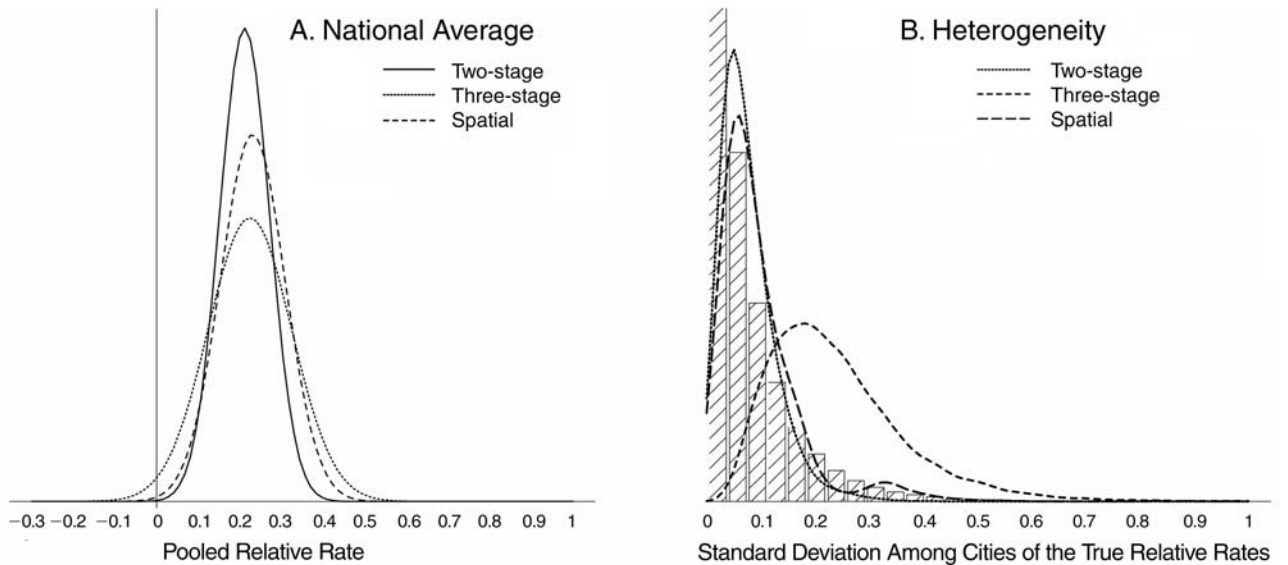
Figure A.1. Marginal posterior distributions of national average estimates of PM_{10} effects on total mortality from nonexternal causes at lag 1. *A.* National average under three models for heterogeneity: 1) two-stage model (baseline approach); 2) three-stage regional model; and 3) spatial correlation model. *B.* Marginal posterior distribution of the standard deviation among cities of the true relative rate (a measure of heterogeneity) under models 1, 2 and 3. The histogram represents the profile likelihood.
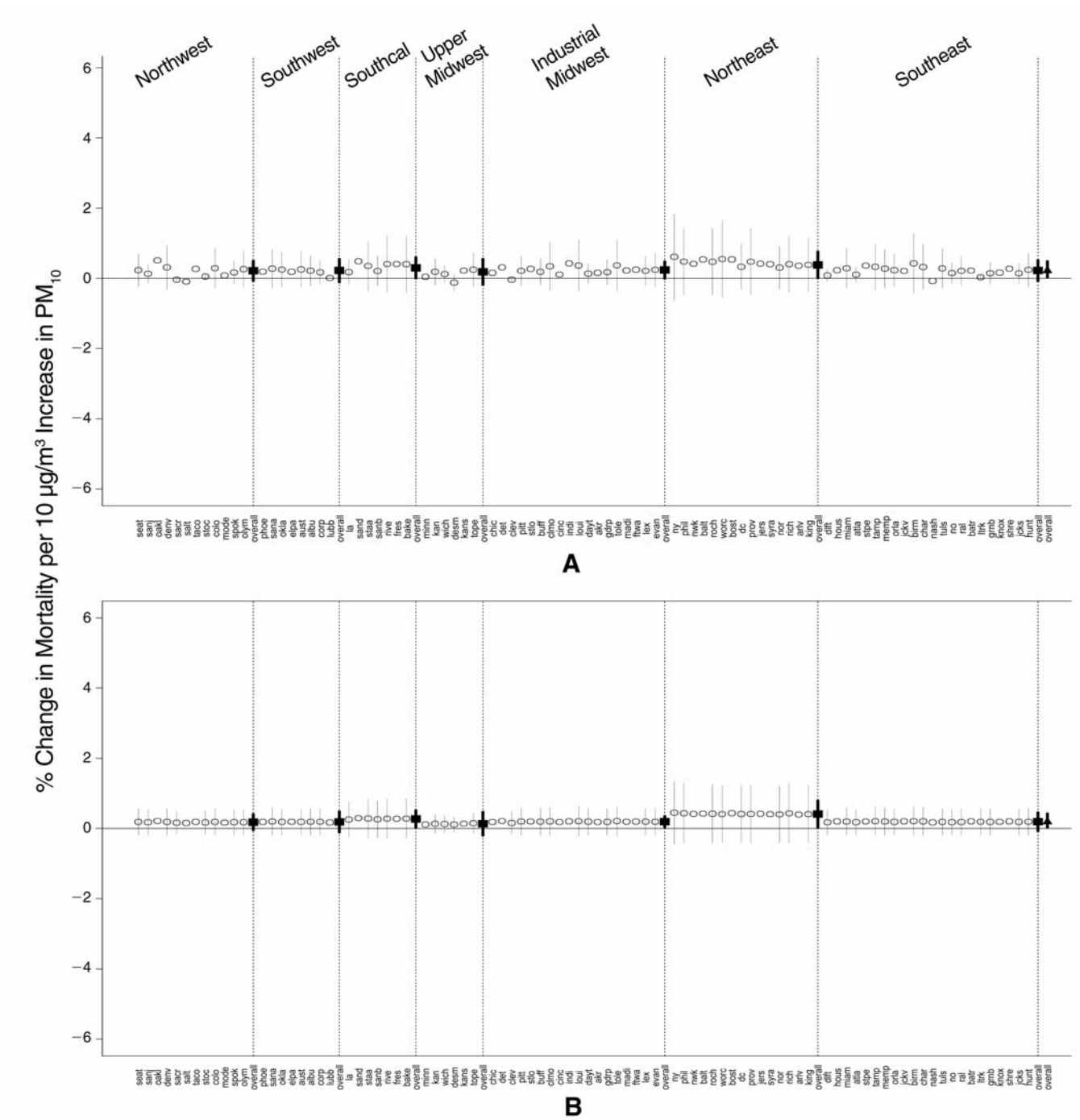
Figure A.2. City-specific posterior means and 95% posterior regions at lag 1 for each of 88 locations under both priors (model A, more heterogeneity within regions; B, noninformative about heterogeneity). Also shown are the posterior estimates and 95% posterior regions at lag 1 for the regional and overall means.

ABOUT THE AUTHORS

Francesca Dominici received her PhD in statistics from the University of Padua in 1997. She is an associate professor of biostatistics at the Johns Hopkins Bloomberg School of Public Health. Her research interests include environmental epidemiology, toxicology and carcinogenicity, metaanalysis of clinical trials, estimating health costs of smoking, risk adjustment models, and estimation of group performance in health services and research.

Aidan McDermott is an assistant scientist in the department of biostatistics at the Johns Hopkins Bloomberg School of Public Health. He holds a PhD from the National University of Ireland, Galway. Dr McDermott's research interests include computational statistics.

Michael Daniels is an associate professor of statistics at the University of Florida in Gainesville. He received his ScD in biostatistics from Harvard University in 1995. His research interests include Bayesian statistics, biostatistics, hierarchical modeling, and estimation of covariance matrices.

Scott L Zeger is professor and chairman of the department of biostatistics at the Johns Hopkins Bloomberg School of Public Health. He received his PhD from Princeton University in 1982. His research interests lie in developing regression analysis for correlated responses.

Jonathan M Samet is professor and chairman of the department of epidemiology at the Johns Hopkins Bloomberg School of Public Health. He holds an MS in epidemiology (1977) from Harvard University; he received his medical degree in 1970 from the University of Rochester School of Medicine and Dentistry. Dr Samet's research interests include the effects of inhaled pollutants, respiratory diseases, cancer, and risk assessment.

ABBREVIATIONS AND OTHER TERMS

CDVRSP	cardiovascular and respiratory diseases
CO	carbon monoxide
df	degree of freedom
gam	function in S-Plus software
GAM	generalized additive model
GLM	generalized linear model
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
NO_2	nitrogen dioxide
O_3	ozone
PM_{10}	particulate matter 10 µm in mass median aerodynamic diameter
SO_2	sulfur dioxide
χ^2	chi-square test

Morbidity and Mortality Among Elderly Residents of Cities with Daily PM Measurements

Joel Schwartz, Antonella Zanobetti, and Thomas Bateson

ABSTRACT

Generalized additive models (GAMs) have been widely used in air pollution epidemiology due to their flexibility in modeling nonlinear factors such as season and weather. Recently several investigators have pointed out problems with implementation of GAMs in some statistical packages. These problems involve lax default convergence criteria and failure to properly estimate standard errors for the parametric terms due to a programming shortcut.

In this report we present a revised analysis of NMMAPS II data. To address the sensitivity of effect size estimates to the choice of convergence criteria, we refit exactly the models originally reported, changing only the convergence criteria. To address the question of sensitivity to the incorrect estimation of standard errors, we present, compared, and applied alternative approaches to nonparametric smoothing that include natural splines, penalized splines, and case-crossover analysis.

We reanalyzed hospital admissions data in 14 US cities using GAM with stricter convergence criteria, natural spline models, and penalized spline models.

We repeated the analyses for mortality data in 10 US cities using GAM with stricter convergence criteria, natural spline models, penalized spline models, and case-crossover analysis.

The overall effect estimate averaged over multiple studies was usually not dramatically different from previously published results and all of the general conclusions previously reported held true in these revised analyses.

INTRODUCTION

GAMs (Hastie and Tibshirani 1990) have been applied in many time-series studies of air pollution and mortality or morbidity. The wide use of GAMs in epidemiologic studies of air pollution is due to flexibility in modeling nonlinear factors such as season and weather (Schwartz 1993, 1994, Katsouyanni et al 2001, Daniels et al 2000).

The *gam* function is available in the S-Plus statistical software (Mathsoft, Seattle WA). A recent report from Dominici and coworkers (2002) indicates that the default convergence criteria used in the S-Plus function for GAM is relatively lax and may not guarantee convergence. This is not a software problem, merely a reminder that investigators need to pay attention to defaults in statistical software because the defaults are not always appropriate for each problem. Independently, Ramsay and colleagues (2003) reported that the S-Plus *gam* function uses a shortcut in estimating covariance of the estimated coefficients that does not properly account for correlation between the exposure variables of interest and the smoothed functions of covariates. This problem may result in biased estimates of the standard errors for the pollution variables.

In view of these reports, we have reestimated our models, allowing for proper convergence and using methods that provide unbiased standard errors. This report provides the results of these revised analyses. To understand the problems, and our attempts to resolve them, we believe it is appropriate to place the problems in context and to discuss what options are available to address the issues.

NONLINEARITY IN EPIDEMIOLOGY

Epidemiologists generally seek to determine whether a variable is associated with health status. When the variable is continuous, we seek to determine how much does a health status change at different levels of the covariate (that is, to estimate the dose-response relation). The two goals are related: badly misspecifying the dose-response

relation for a continuous predictor can lead to false inferences about the strength of the association.

Dose-response relations need not be linear. For example, increased blood-lead levels have little impact on erythrocyte protoporphyrin concentration at low doses, but the impact rises nonlinearly at higher doses. In this example, toxicokinetic modeling can provide an indication of the functional form of the relation (Marcus and Schwartz 1987). In most instances, a physiologically based model is not available and empirical approaches must be used. To test hypotheses, it is common to see continuous variables categorized with, for example, high or zero alcohol consumption contrasted with moderate intake.

If the variable suspected of a nonlinear association is not the hypothesis variable, but a control variable, this misspecification of the dose response (it assumes no change with increasing dose within category) risks allowing residual confounding.

Piecewise constant fits (that is, indicator variables for categories) are an example of piecewise polynomial fits (that is, separate polynomials for each range of the predictor variable). A piecewise linear fit within each category is a natural extension that tries to capture some of the variability in outcome with variability in exposure within intervals of exposure. Piecewise linear fits are examples of **regression splines**, which include higher order polynomials.

It has long been noted with piecewise constant fits that the shape of the dose-response curve (created by plotting the fit) is sensitive to the cut points chosen for the different categories. This has lead to the widespread use of quartiles or quintiles to define the cut points (called knots in the spline context). This categorization removes the choice of cut point from the discretion of the investigator and prevents misrepresentation, but it does not eliminate the sensitivity. Higher order piecewise polynomials share this sensitivity to the location (and number) of knot points. Sensitivity can be reduced by increasing the number of categories. Using too many categories, however, results in curves that are too wiggly. This wiggliness also presents a problem for model building. Using automatic knot point selection on quantiles means that the location of all knots changes when one degree of freedom (*df*) is added to a model. Consequently, the curves jump about quite a bit. In addition, changing from an odd to an even number of knots eliminates the knot at the median. Possible problems from using this approach are illustrated in Figure 1. Two estimated dose-response curves between pneumonia admissions in Detroit and temperature are dramatically different when 3 *df* rather than 4 *df* is used.

In the regression spline field, a common approach to this problem is to use stepwise regression techniques to
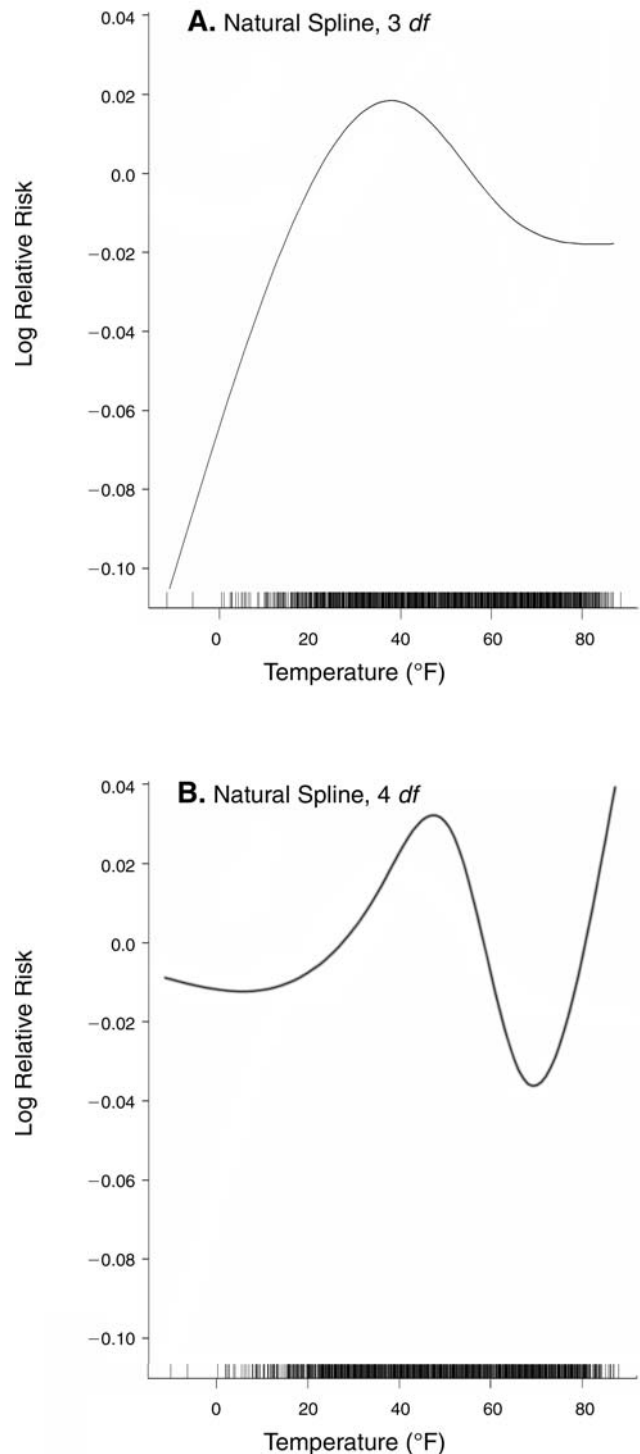
Figure 1. Detroit. Relation between pneumonia hospital admissions and temperature using natural spline with 3 *df* (*A*) and with 4 *df* (*B*).

choose the location of the knots (the number of knots can also be addressed this way). To date, such methods have not been applied to reanalyze air pollution data.

One alternative approach advocated by Eilers and Marx (1996) is called **penalized splines**. The analyst chooses a fairly large number of categories (thereby reducing the sensitivity to knot location) and then constrains the coefficients within each interval by imposing a penalty for the wiggliness of the resulting curve. Instead of minimizing the sum of squares error, or more generally the log likelihood, one minimizes the penalized likelihood. This penalized likelihood is the sum of the usual log likelihood and a parameter multiplied by the sum of the squares of the slope changes. Specifically, consider a model of the form:

$$Y_i = \beta_0 + \beta_1 x_1 + \sum_{k=1}^{K} u_k (x_1 - K_k)_+ + e_i$$

where K is the number of boundary points (knots) between intervals of x, K_k are the locations of those points, and $(x_1 - K_k)_+$ is defined as:

$$\begin{cases} x_i - K_k, x_i > K_k \\ 0, otherwise \end{cases}$$

So far, this is simply a piecewise linear fit. If we now fit this using a penalized likelihood, we constrain the u_k at each knot point, effectively reducing the degrees of freedom of the piecewise fit. This can be done by the following constraint:

$$\sum_{k=1}^{K} u_k^2 < C$$

In the Gaussian case, instead of minimizing the sum of squared errors, we would add a penalty term

$$\lambda^2 u^T D u$$

where D is a suitably defined design matrix. The extension to cubic or other splines is immediate.

The larger the parameter λ (called the *smoothing parameter*) that multiplies the penalty term, the smaller the changes in slope that will minimize this penalized likelihood, and hence the smoother the curve. The degree of smoothness (and hence the value of the parameter) can be estimated by using generalized cross-validation or other methods. One advantage of this approach is that as one increases the degree of freedom for temperature, the locations of the knots do not change. Because the number of knots is larger than when using regression splines with the same degree of freedom, sensitivity to knot location decreases.

Software to use penalized splines to fit multiple covariates to a health outcome, including logistic and Poisson regressions, has been developed by Wood (2000) and is available in R (*www.cran.r-project.org/*). This software uses cubic or thin plate splines.

Brumback and coworkers (1999) have pointed out that this can all be expressed in terms of a mixed model. Because mixed models are increasingly used for longitudinal data, geographical data, stratified sampling, and multilevel models, mixed models offer a unified framework to address many issues simultaneously. Specifically, if one assumes that the u_k are not fixed, but random effects, then:

$$u_k \sim N(0, \Omega)$$

We have a mixed model with random slopes u_k. Because they are declared random, the mixed model machinery will constrain the estimates of the u_k to vary less than if they were treated as fixed effects. This constraint reduces the degrees of freedom, just as the penalized likelihood formalism. The amount of shrinkage can be prespecified ("I want a 3 *df* curve…") or by restricted maximum likelihood (REML). By using REML, one allows the data themselves to determine the appropriate amount of constraint. The mixed model programs in most statistical packages will estimate those random slopes, allowing the user to rely on standard mixed model theory for model fitting and inference.

One feature of this mixed model formalism is that it allows more than one covariate to have a nonlinear relation to the outcome, while also allowing the more usual mixed model assessment of heterogeneity. Coull and coworkers (2001) recently used such an approach to show heterogeneity in response to air pollution in a panel study, while controlling for a nonlinear dependence on temperature and season. This approach can be extended to generalized linear mixed models (Ruppert et al 2003), allowing logistic and Poisson models as well. While, for simplicity, we have illustrated these approaches for linear splines, the extension to quadratic or cubic splines is immediate.

Another approach is to use GAM (Hastie and Tibshirani 1990). These models have been applied in many time-series studies of air pollution and mortality or morbidity. The wide use of GAM is due to its flexibility in control of the nonlinear effects of confounding variables such as trend, seasonality, and weather variables. In most of those studies, air pollution has been treated as a linear term, and the nonlinear capabilities of GAM have been used for covariate control, the focus of this report.

GAM is an extension of generalized linear modeling (GLM) (McCullagh and Nelder 1989) that allows modeling of nonlinear effects by using nonparametric smoothing functions. Nonparametric smoothers can be viewed as

extensions of moving averages. In moving averages, the estimated effect Y_i at a point X_i is taken to be a weighted average of all of the Y_i in a neighborhood about X_i. The neighborhood is called a window, and the weights generally decline to zero between X_i and either edge of the neighborhood. These are called kernel smoothers, where the weights are the kernel. The size of the window determines the degree of smoothing, wider windows producing smoother curves. The extensions include the use of weighted regressions, locally weighted smoothing function (LOESS) instead of weighted averages (Cleveland and Devlin 1988). Smoothing splines are derived differently, using penalized likelihood but with a knot at each data point, unlike the penalized splines.

One important difference between the GAM of Hastie and Tibshirani and the penalized or regression spline models is the method of estimating multiple curves. GAM uses the backfitting algorithm to sequentially smooth against one covariate at a time, iterating until convergence. In contrast, regression splines and penalized splines (using either the ridge regression or the mixed model formulation) estimate all smooth terms in a single step.

The *gam* function available in S-Plus software was used in the NMMAPS project to estimate associations between air pollution and hospital admissions and to estimate the distributed lag between air pollution and daily deaths. The GAM was among the first flexible extensions of the GLM and has been widely used in air pollution epidemiology.

One alternative approach to covariate control is matching. If we match cases and controls by a covariate (say season or weather), we do not have to worry about nonlinearity in modeling their association with outcome. This can be done within the context of a **case-crossover analysis**. The case-crossover design, introduced to epidemiologists by Maclure (1991), offers an attractive method to investigate the acute effects of an exposure. For example, the method has been used to investigate triggers of myocardial infarction (Mittleman et al 1993). In recent years, it has been applied to analysis of the acute effects of environmental exposures, especially air pollution (Lee and Schwartz 1999; Neas et al 1999; Sunyer et al 2000; Levy et al 2001). In the case-crossover approach, a case-control study is conducted whereby each person who had an event is matched with herself on a nearby time period during which she did not have the event. The subject's characteristics and exposures at the time of the case event are compared with those of a control period in which the event did not occur. Each risk set consists of one individual as that individual crosses over between different exposure levels in the interval between the two time periods. These matched pairs may be analyzed

using conditional logistic regression. Multiple control periods may be used.

The data are then analyzed as a matched case-control study. Applied to the association of air pollution with risk of death, the approach has several advantages. First, it clarifies a key feature of the study of acute response to air pollution. Because in this analysis each subject serves as her own control, the use of a nearby day as the control period means that all covariates that change slowly over time (such as smoking history, age, body mass index, usual diet, diabetes mellitus) are controlled for by matching.

The case-crossover design controls for seasonal variation, time trends, and slowly-varying covariates because the case and control periods in each risk set are separated by a relatively small time interval. Bateson and Schwartz (1999, 2001) demonstrated that by choosing control days close to event days, even very strong confounding of exposure by seasonal patterns could be controlled by design in the case-control approach. Also, because the analysis is of matched strata of days for each individual, it is straightforward to combine events from multiple locations in a single analysis. The difference in seasonal patterns from city to city has prevented this approach in multicity studies using Poisson regression. This makes the approach an attractive alternative to the Poisson models.

While Bateson and Schwartz (2001) have shown that the power to detect small effects is lower in the case-crossover approach, this is less of a concern in a large multicity study. Concern has been expressed that the symmetric bidirectional sampling of controls proposed by Bateson and Schwartz (1999) violates the study base principal and does not have a proper likelihood, resulting in some bias in the estimated coefficients (Navidi and Weinhandl 2002). However, Bateson and Schwartz have shown that this is equivalent to selection bias. That is, each day does not contribute equally to cases and controls. For example, the days before start of the time series can contribute control exposure, but not case exposure. They showed that this bias could be estimated and subtracted, and they provide details in their 2001 paper.

To further address this, we have simulated 5000 time series, each with several different patterns of confounding in both outcome and exposure: confounding by

- random noise added to both exposure and outcome,
- linear time trend,
- 365-day cosine term,
- combination of a strong linear trend with the cosine pattern,
- highly structured time series based on a smoothed function of pneumonia hospital admissions, and

- actual pattern of particulate matter less than 10 μm in aerodynamic diameter (PM$_{10}$) in Denver, Colorado.

The last pattern was used by Navidi and Weinhandl (2002) to illustrate the potential for bias in the unadjusted case-crossover approach using our sampling scheme. The latter represents a series with considerable structure at multiple wavelengths. Using the simulations, we present below the mean of the estimated coefficients (after bias correction) and the 95% coverage probabilities of those estimates. The true log relative risk is 0.100 in all cases. Controls were taken at ± 6 and 7 days (Table 1).

As we have demonstrated, the method appears unbiased with coverage probabilities close to 95%. Hence the standard errors are also unbiased. We have compared performance of the time-stratified sampling scheme of Levy and colleagues (2001). The two approaches give comparable results. We replicated the large bias using the naïve symmetric bidirectional (SBI) control sampling that Navidi had reported; however, as discussed previously in this report, that bias was eliminated by our correction procedure. Hence, case-crossover analyses can be applied to control for season using matching, which avoids the issues involved in the GAM problems. In principal, we can also choose controls that are matched on weather as well as season, thereby controlling both of the realistic confounders in air pollution time series. We are doing further work using that approach.

In this report, we present results of the revised analyses of NMMAPS morbidity and distributed lag data. In those analyses, we have at a minimum used stringent convergence criteria in reestimating effects with GAM. In addition, we have used at least one of the alternative techniques described previously to refit our models using the same degree of freedom as the original GAM but with correct estimation of the standard errors.

DATA

In the NMMAPS study (Samet et al 2000), we evaluated the association between hospital admissions of persons aged 65 and older in 14 US cities and PM$_{10}$ for the following illnesses, as defined by primary discharge diagnosis:

- cardiovascular disease (CVD) (*International Classification of Diseases*, 9th Revision [ICD 9], 390–429),
- pneumonia (ICD 9, 480–487), and
- chronic obstructive pulmonary disease (COPD) (ICD 9, 490–492, and 494–496).

Table 1. Results of Symmetric Bidirectional Case-Crossover Analyses of Simulated Time-Series Data in Presence of Time-Varying Confounding With and Without Adjustment for Selection Bias[a]

Patterns in Simulated Data		Naïve Results		Results Adjusted for Selection Bias		Time-Stratified Results	
Time Trend in Exposure[b]	Time Trend in Outcome	Mean Coefficient	95% Coverage Probability[c]	Mean Coefficient	95% Coverage Probability[c]	Mean Coefficient	95% Coverage Probability[c]
Random[d]	Random[d]	0.10043	93.68	0.10018	94.66	0.0995	94.84
Linear[e]	Linear[e]	0.09908	93.36	0.09933	94.74	0.0996	94.80
COS1[f]	COS1[f]	0.10009	93.82	0.10124	94.92	0.1033	94.52
COS1×Linear[g]	COS1×Linear[g]	0.09969	94.08	0.10122	95.28	0.1084	94.48
Pneumonia	Pneumonia	0.10044	94.44	0.10006	95.18	0.1019	95.10
Denver PM$_{10}$	COS1×Linear[g]	0.15964	77.6	0.10041	94.32	0.0915	94.92

[a] The simulated ln(RR) was 0.1000. Results are based on 5000 iterations.

[b] The Confounding Time Trend was used to generate the exposure data as well being used to simulate an omitted covariate that affected the daily event occurrence.

[c] Coverage probabilities are the percentage of 95% confidence intervals that contain the true effect.

[d] Random means identical random Gaussian noise was added to both exposure and outcome.

[e] Linear refers to linear time trend in both exposure and outcome with an amplitude of 60% of the mean number of events, and a frequency of 1 cycle per year. This is very large seasonal cycle compared to most mortality or morbidity data.

[f] COS1 refers to cosine term in both exposure and outcome with an amplitude of 60% of the mean number of events, and a frequency of 1 cycle per year. This is a very large seasonal cycle compared to most morbidity and mortality data.

[g] COS1×Linear refers to a combination of the linear time trend times the seasonal cosine trend.

For a detailed description of the health data, refer to the NMMAPS report pages 28 to 47 (Samet et al 2000). In the same report, we also examined the distributed lag between PM_{10} exposure and daily deaths in 10 US cities with daily PM_{10} measurements. That analysis is more fully described in Appendix B of the original NMMAPS report (Samet et al 2000).

METHODS

GAM WITH STRICTER CONVERGENCE CRITERIA

To address the sensitivity of the effect estimates to the choice of convergence criteria and maximum iterations (for the main and backfitting steps) in GAM, we adopted the strategy of refitting *exactly the model originally reported*, changing only the convergence criteria. This was done by rerunning the original script files. The details of the original models for each outcome and each city are given in the original NMMAPS report (Samet et al 2000) in Appendix D, Tables D.2 through D.4. The stricter convergence criteria were the ones recommended by Dominici and coworkers (2002): `maxit = 1000,bf.maxit = 1000; epsilon = 10e-15,bf.epsilon = 10e-15.`

Use of the stricter criteria addresses the possibility that models may not have converged in the original analyses but not the issue of standard error estimation. This second issue is much more of a problem for single-city studies than for multicity studies such as NMMAPS or Air Pollution and Health: A European Approach (APHEA). In the latter studies, investigators do not simply rely on the estimated variance of regression coefficients for air pollution in a single city. Rather, they have the empirical distribution of coefficients for air pollution across multiple cities. If the within-city estimates of variance are too small, the observed variation in results between cities will be larger than expected from the within-city variances. This results in the use of a random variance component to estimate the difference between the within-city and between-city estimates of variance of the coefficient. The standard error of the combined analysis is determined by the overall variance of the coefficients. Hence, underestimating the variance of coefficients within individual cities would not be expected to have much effect on variance of the overall mean estimate. We report empirical results here to confirm this point.

ALTERNATIVES TO GAM

To address the question of sensitivity to incorrect estimation of standard errors within city is more difficult.

There is no conceptual difficulty in fixing the problem for the GAM software, and a fix is expected to be available shortly. In the interim, any change necessarily must change not only the way standard errors are estimated, but also how to control for season and weather. Hence, effect estimates will change as well as confidence intervals (CIs). Because of time limitations, we have not been able to repeat all analyses all ways. We have repeated all of our analyses using GAM, with the stricter convergence criteria, using penalized spline and natural spline models with the same degrees of freedom. Because we did not need a mixed model for any part of our analysis, we used the generalized additive penalized spline program in R, which is more straightforward to implement. We used thin plate splines as the basis functions for regression splines in the penalized spline model. For the natural spline models, noninteger degrees of freedom were rounded to the nearest whole number. Autocorrelation of the residuals of the new models was still so-called white noise. For some cities, autoregressive terms were added as in our original models.

We have also applied the case-crossover approach to some mortality data to illustrate the technique. All of the results are reported below in summary, with more details in the appendices.

CHANGE OF METHOD TO COMBINE RESULTS ACROSS CITIES

In our original report of these results, we used the method of moments to estimate the random city effect (that is, the measure of heterogeneity). This is an approximate method. Since then, we have used iterative maximum likelihood estimation of the random city effect variance, using the method of Berkey and coworkers (1998). We have used this more accurate method to combine the results across cities and to obtain the χ^2 statistic for the heterogeneity test. A description of this method follows.

To combine the coefficients across cities, we used the city-specific estimates and the variance-covariance matrix of the parameter vector. This was achieved by fitting a metaregression model based on the method described by Berkey and colleagues (1998). More specifically, models for metaregression are of the form:

$$\beta^c = X^c \alpha + \delta^c + \varepsilon^c$$

where β^c is the (14×1) vector of estimates in each city; X^c is the identity matrix plus any factors that predict the heterogeneity across cities; α is the vector of regression coefficients to be estimated; δ^c is a vector of random effects associated with city c and ε^c (assumed independent from δ^c) is the vector of sampling errors within each city.

The matrix $cov(\delta^c) = D$ (needed to be estimated) represents the between-cities covariances of regression coefficients that is unexplained by the sampling errors.

It is assumed that:

$$\delta^c \sim N(0,D)$$

$$\varepsilon^c \sim N(0,S^c)$$

$$\beta^c \sim N(X^c\alpha,D+S^c)$$

where S^c is the estimated variance-covariance matrix in city c. When $D \approx 0$, we get the corresponding fixed effects estimates, but when $D \neq 0$, we get the random-effects estimates.

The method described by Berkey and colleagues (1998) is applied to estimate the model parameters. More specifically, the iterative generalized least squares method was applied to estimate model parameters. That is,

$$\hat{\alpha} = (\sum_c X^{c^T}V^{c^{-1}}X^c)^{-1}(\sum_c X^{c^T}V^{c^{-1}}\beta^c) \quad \text{with}$$

$$Var(\hat{\alpha}) = (\sum_c X^{c^T}V^{c^{-1}}X^c)^{-1}$$

where $V^c = S^c$ for the fixed effects estimates and $V^c = S^c+D$ for the random effects estimates. The parameters of the between-cities covariance matrix D are estimated by maximum likelihood (Berkey et al 1995). The overall test for heterogeneity is

$$Q = \sum_c (\hat{\beta}^c - X^c\hat{\alpha})^T \hat{V}_{FE}^{c^{-1}}(\hat{\beta}^c - X^c\hat{\alpha})^T$$

and under the null hypothesis of homogeneity ($D \approx 0$) follows the χ^2 distribution with degrees of freedom equal to the number of cities minus the number of covariates (1 in our case), which gives us a total of 13 *df*. Because of the limited power of this test, the results we report are all from random effects metaanalyses, regardless of whether there was significant heterogeneity.

RESULTS

MAIN RESULTS COMPARING ANALYTIC METHODS

We here include the results of various reanalyzes we presented originally in the NMMAPS report. In the main body of this report, we show summary results across all cities; city-specific analyses are included in Appendix A.

Hospital Admissions

Table 2 shows the combined mean estimates for constrained lag (one-day mean, two-day mean, quadratic distributed lag) and unconstrained distributed lag models for the percentage change in hospital admissions for a 10-$\mu g/m^3$ increase in PM_{10}. These results are for the original GAM the GAM with stricter convergence criteria, and the natural spline models. Table 2 corresponds to Table 14 in the original report. In Table 2, we present the combined results from the random effect estimates, while in Table 14 of the original report fixed effect estimates are reported (Samet et al 2000, p 37).

The differences between the combined estimates for the two-day mean of PM_{10} (obtained with GAM with original and stricter criteria) were: reductions of 7.6% for hospital admissions for CVD, 8.9% for COPD hospital admissions, and 9.6% for pneumonia hospital admissions. These differences are all smaller than the corresponding drop of 34% observed in the NMMAPS mortality analysis for 90 cities. (Percentages were computed as a percentage change in the regression coefficients from the combined analysis.) This difference is likely due to the greater degrees of freedom included in the models applied to analyze mortality data, which would have led to the original mortality estimates being further from convergence than those obtained with our models with fewer degrees of freedom applied to analyze the morbidity data. The results using natural splines were decreased by 11.2% for CVD admissions, by 28.5% for COPD admissions, and by 70% for pneumonia admissions compared to the original results. For comparison, the NMMAPS mortality effect estimates dropped by 48.8% between the original GAM fits and the natural spline fits. Thus, with the exception of pneumonia and natural splines, the reduction in effect sizes was generally smaller for the NMMAPS morbidity analyses.

For distributed lag models, we found that for CVD the estimates with the stricter convergence criteria increased by 4% and 5% (respectively for quadratic and unconstrained distributed lag) and that, using natural spline, they increased by 1% and 5%. The latter results are interesting because Le Tertre and colleagues have also reported that a modest increase in effect estimates for CVD hospital admissions in the APHEA study when natural spline models were applied (Le Tertre et al, elsewhere in this Special Report).

For COPD, the GAM fits with stricter convergence criteria resulted in reductions of 14% and 12%, respectively, while the reductions using natural splines were 28% and 26%. The reductions in distributed lag models for COPD

Table 2. Percentage of Change in Hospital Admission for Specific Diagnoses per 10 µg/m^3 Increase in PM$_{10}$ in 14 Cities

	CVD		COPD		Pneumonia	
	% Change	95% CI	% Change	95% CI	% Change	95% CI
Original GAM with LOESS—Random Effect[a]						
Constrained lag models						
One-day mean	1.02	0.69, 1.35	1.42	0.35, 2.51	1.62	1.05, 2.18
Previous-day mean	0.68	0.54, 0.81	1.18	−0.32, 2.70	1.31	1.03, 1.58
Two-day mean	1.17	1.01, 1.33	1.81	1.16, 2.47	1.90	1.44, 2.37
PM$_{10}$ < 50 µg/m^3 (two-day mean)	1.45	1.12, 1.78	2.60	1.40, 3.81	2.46	1.16, 3.78
Quadratic distributed lag	1.05	0.67, 1.43	2.94	0.21, 5.74	1.87	0.72, 3.04
Unconstrained distributed lag	1.06	0.67, 1.46	2.88	0.22, 5.61	2.07	0.93, 3.23
GAM with LOESS and Stricter Convergence Criteria[b]						
Constrained lag models						
Two-day mean	0.99	0.79, 1.19	1.71	0.95, 2.48	1.71	1.16, 2.26
Quadratic distributed lag	1.09	0.81, 1.38	2.53	1.20, 3.88	1.47	0.86, 2.09
Unconstrained distributed lag	1.12	0.84, 1.40	2.53	1.21, 3.87	1.62	0.95, 2.29
Natural Spline Model[b]						
Constrained lag models						
One-day mean	1.01	0.76, 1.26	1.09	0.48, 1.70	0.78	0.45, 1.11
Previous-day mean	0.52	0.29, 0.76	0.99	0.25, 1.73	0.22	−0.45, 0.40
Two-day mean	0.96	0.71, 1.20	1.32	0.56, 2.08	0.57	0.04, 1.10
PM$_{10}$ < 50 µg/m^3 (two-day mean)	1.32	0.77, 1.87	2.21	1.02, 3.41	1.06	0.06, 2.07
Quadratic distributed lag	1.11	0.79, 1.44	2.11	0.63, 3.61	0.01	−0.64, 0.67
Unconstrained distributed lag	1.06	0.72, 1.40	2.10	0.63, 3.59	−0.02	−0.65, 0.61

[a] Methods of moments.

[b] Combined effect estimates obtained using maximum likelihood estimation.

were quite similar to those for the 2-day mean PM$_{10}$. For pneumonia, we observed a reduction of 21% and 22% when using the stricter convergence criteria, while the results for distributed lag models using natural spline became null. This result deserves some discussion.

Pneumonia is the outcome measure in our study that shows the greatest seasonality. Pneumonia cases peak in the winter, but in many of the cities important PM$_{10}$ peaks occur in the summer. Hence, control of season is critical. Notably, when our analysis was restricted to days with PM$_{10}$ below 50 µg/m^3, which included most days but omitted the high-PM summer days, PM$_{10}$ was significantly associated with daily pneumonia admissions in the natural spline analysis. While the effect size was halved for pneumonia, it barely changed for CVD and COPD admissions compared to the old results. These results suggest that the pneumonia effect estimates were sensitive to peaks in exposure during a season when pneumonia is rare. Interestingly, for pneumonia there was evidence of confounding by other

pollutants; control for ozone, another summer peaking pollutant, increased the effect of PM$_{10}$.

To further examine this issue, we fit a model for Pittsburgh (one of the cities where the effect estimates became negative) using a natural spline of PM$_{10}$ (mean of lag 0 and lag 1). Figure 2 shows the results of that model. At low to moderate PM$_{10}$ values, hospital admissions for pneumonia increased with increasing concentrations, but they then fell sharply when PM$_{10}$ levels were above 30 µg/m^3. The highest days were concentrated in July and August. This finding suggests that further examination of pneumonia admissions restricted to the winter season, when the bulk of them occur, may be of interest. Time constraints have prevented us from doing so to date.

Table 3 shows the summary results of repeating analyses for the three outcomes using penalized splines in R. City-specific results are presented in Appendix A. The results using penalized smoothing splines were intermediate between the GAM and natural spline results. In particular,
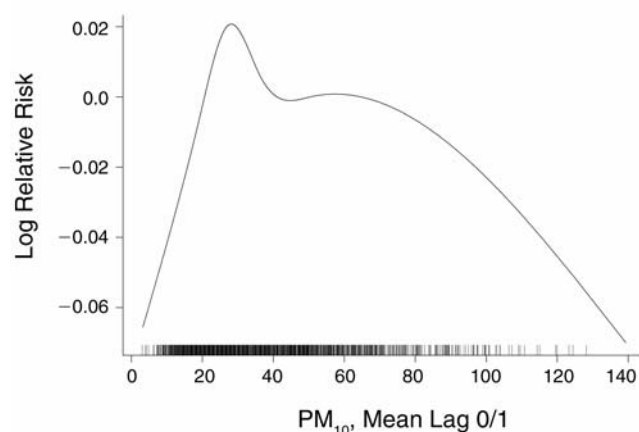
Figure 2. Pittsburgh. Relation between pneumonia hospital admissions and PM_{10} at mean lag 0/1 using natural spline with 5 *df*. At low to moderate PM_{10} values, hospital admissions for pneumonia are shown to increase with increasing concentrations but then to fall sharply when PM_{10} levels are above 30 µg/m^3.

using penalized splines, a significant effect continued for pneumonia using a distributed lag model. We believe this warrants further study. Given that GAM, penalized splines, and regression splines gave similar results for CVD and COPD, but that regression splines produced dramatically different results for pneumonia admissions, we believe that the natural spline model may be less stable, possibly because of the sensitivity to knot location that penalized splines are designed to remove. We also note that Dominici and colleagues (2002) observed greater noise in the estimates from natural splines than from GAMs as they varied the degree of freedom in their models. The penalized spline approach may offer a good compromise between regression and smoothing splines by giving appropriate standard errors and avoiding backfitting, while preserving the greater flexibility of smoothing.

Table 4 shows the test for heterogeneity for models using natural splines. These results are similar to those presented previously (Samet et al 2000, Table 15, p 38). χ^2 sta-

Table 3. Percentage of Change in Hospital Admission for CVD, COPD, and Pneumonia per 10 µg/m^3 Increase in PM_{10} in 14 Cities[a]

	Two-Day Mean		Quadratic Distributed Lag		Unconstrained Distributed Lag	
	% Change	95% CI	% Change	95% CI	% Change	95% CI
CVD	1.00	0.80, 1.19	1.09	0.81, 1.37	1.12	0.83, 1.41
COPD	1.56	0.86, 2.28	2.39	1.13, 3.67	2.54	1.20, 3.89
Pneumonia	1.23	0.49, 1.98	0.64	0.03, 1.25	0.80	0.14, 1.46

[a] Penalized spline model. Combined effect using maximum likelihood estimation.

Table 4. Heterogeneity Test for City-Specific Estimates for Specific Diagnoses[a]

	PM_{10} Lag 0	PM_{10} Lag 1	PM_{10} Lag 0/1	Quadratic Distributed Lag	Unrestricted Distributed Lag	PM_{10} Lag 0/1 (<50 µg/m^3)
CVD						
Heterogeneity χ^2	23.21	23.1	21.04	18.9	17.78	30.49
df	13	13	13	13	13	13
P value	0.04	0.04	0.07	0.13	0.17	0
COPD						
Heterogeneity χ^2	16.19	23.11	20.31	32.29	32.47	16.07
df	13	13	13	13	13	13
P value	0.24	0.04	0.09	0	0	0.25
Pneumonia						
Heterogeneity χ^2	11.17	18.49	24.02	17.66	18.39	22.97
df	13	13	13	13	13	13
P value	0.6	0.14	0.03	0.17	0.14	0.04

[a] Natural spline model.

tistics, degree of freedom, and *P* values from the iterative maximum likelihood estimation of the random city effect variance program were used to obtain our combined-effect estimates. We do not present the decomposition among total, within, and random variance, but this can be calculated from the presented results.

Mortality

The original NMMAPS report contained our analysis of 10 US cities with daily PM_{10} monitoring to explore the distributed lag between air pollution and daily deaths (Samet et al 2000). We have similarly repeated these analyses using the stricter convergence criteria, penalized splines, and natural splines. Table 5 shows the estimated overall effect of PM_{10} looking at lags out to 5 days. The effect size estimates are reduced compared to the original report (Samet et al 2000 Appendix B, pp 54–61, and Table B.3, p 59). Once again, however, these estimates are noticeably larger than those obtained using single-day exposures, indicating that effects of a given concentration of PM_{10} on one day occur both on that same day and for several days afterward. This observation has been confirmed in many other studies. In this instance, the penalized spline results differed little from the natural spline results.

We also used the case-crossover analysis with these 10 cities (Appendix B of this report). Matched strata constructed for each subject consisted of the event day (day of death) and 18 matched control days. These days were chosen to be days 7 to 15 before the event day and days 7 to 15 after the event day. We chose control days far from the event day to avoid serial correlation in the pollution and mortality data. Control days were symmetrically chosen about the event day because Bateson and Schwartz (1999) demonstrated that symmetric control days were needed to control for long-term time trends (if present). Navidi (1998) pointed out that bidirectional sampling is needed to avoid some biases in the case-control study and does not present any conceptual difficulties as long as death of the subject does not affect the air pollution concentrations. Regression splines were used to control for temperature and humidity on the day of death and the day before death. We used the bias correction procedure of Bateson and Schwartz (2001), which performs well with respect both to bias and coverage probabilities. For ease of comparison with the main NMMAPS mortality analysis, we used PM_{10} concentrations the day before death and the day before each of the control days as the exposure index. These results are shown in Table B.3.

It is interesting to compare these results with the NMMAPS results for mortality in 90 cities. Rather than try to find the optimal degree of freedom per year for seasonal

Table 5. Daily Deaths in 10 US Cities: Percentage of Change per 10 μg/m³ Increase in PM_{10}

	% Change	95% CI
Original GAM Results		
Quadratic distributed lag	1.4%	1.2, 1.7
Unconstrained distributed lag	1.3%	1.0, 1.5
GAM with Stricter Convergence Criteria		
Quadratic distributed lag	1.2%	0.9, 1.5
Unconstrained distributed lag	1.1%	0.8, 1.4
Natural Spline Model		
Quadratic distributed lag	1.0%	0.6, 1.4
Unconstrained distributed lag	1.0%	0.7, 1.3
Penalized Spline Model		
Quadratic distributed lag	1.0%	0.8, 1.3
Unconstrained distributed lag	1.0%	0.8, 1.3

control, the NMMAPS mortality study presented results using a range of values. This is still true in the reanalyses using natural splines. The magnitude of association using case crossover is larger than that in NMMAPS natural spline models for 90 cities, which use 7 *df*/year to control for season (0.35% vs 0.21%). This larger association could be due to sampling variability, to the choice of cities studied, or to the difference in methods. If it is due to the latter, the question is most likely degree of freedom rather than use of natural splines per se because any sensitivity to knot location should average out in a large multicity study, such as NMMAPS. Most other studies of air pollution and daily deaths have used 3 to 4 *df*/year (Goldberg et al 2001; Katsouyanni et al 2001). The midpoint of this range (3.5 *df*/year) is one of the choices used in the NMMAPS study, and results of that analysis are almost identical to results of the case-crossover analysis.

Simulation studies previously conducted (Bateson and Schwartz 1999, 2001) and those noted earlier in this report have demonstrated that the control sampling strategy used in this analysis is adequate to control for strong seasonal confounding. Thus it follows that fewer degrees of freedom may be more appropriate for spline models.

There is really no correct choice for how to control for season and weather. This ill-posed problem cannot be answered from the data alone. Fortunately, we have additional information on which to rely. The reason we control for season is that we believe covariates have been omitted that will be accidentally correlated with air pollution variations over long timescales but which are unlikely to be correlated over short timescales. For example, both air

pollution and smoking rates have fallen over time. Failure to control for time trends could therefore result in confounding. We think it unlikely that this would explain one-week excursions in mortality rates that occur coincidentally with an air pollution episode. Similarly, many risk factors vary by season, as does air pollution. Hence we use smooth functions of time to filter out fluctuations with periods greater than a certain timescale. We cannot ask the data to tell us the appropriate choice. What we can do is look at the fitted curves, see what timescales they are removing, and choose degrees of freedom that are consistent with our intuition (and data) about the likelihood of confounding.

Examination of the partial autocorrelation function of the residuals is also helpful. The impact of omitted covariates is in the residuals of the model. If those covariates have no pattern of fluctuation over time, they are unlikely to correlate with air pollution unless one can identify a physical connection between them and air pollution. Hence, the presence of serial correlation in the residuals indicated the potential for confounding. Reducing the residuals to white noise is therefore a good thing, but it cannot guarantee freedom from confounding. An omitted confounder may have too small an effect on mortality to significantly affect the residuals but still confound air pollution.

Fortunately, the obvious confounders of air pollution are related to weather, and they are available and measured in our studies. Smoking and dietary habits are unlikely to covary with air pollution on timescales briefer than seasons, and we can bring this information to our model choice. Similarly, we believe extremely cold or hot temperatures are bad for health. It is physiologically not meaningful, however, to postulate that the risk of death will, for example, rise from 60 to 63°F, fall from 63 to 67°F, and then resume rising. Temperature effects may be non-linear, but they should be quite smooth. This can also guide our modeling efforts.

 The problems associated with overcontrol of season can be seen by looking at two examples. Of course, one may wonder whether there can be any bias from overcontrol. Normally, one would expect increases in standard errors, and noisy estimates in individual cities, but no bias. Time-series analysis is somewhat different, however. The effects of exposure to an environmental variable (pollution or weather) are usually distributed over time. Because pollution and weather are serially correlated, the use of a single day's exposure (as in the 90 cities of NMMAPS) or the mean of 2 days' exposure (as in APHEA, Katsouyanni et al 2001) captures not merely the effect of exposure on that time scale, but part of the effect that is manifested with longer lags. In this sense, the estimate is confounded, reflecting an overestimate of the effect of exposure on one

day, for example, but an underestimate of the cumulative effect. Increased filtering of the time series will eliminate this bias and give one the instantaneous effect of exposure. This is why in classical time series, where the lag pattern has important policy implications (eg, the lag between interest rate cuts and increases in gross domestic product), prewhitening is used. The effect over time is examined using distributed lag models.

The main public policy question this study evaluates is the magnitude of short-term effects of air pollution on mortality. As noted, the estimated effect of exposure after one day can be an overestimate of the effect with that lag but an underestimate of the overall. Increased filtering of the time series beyond that which is necessary to control for season can lead to further underestimates. Moreover, because the filtering can remove patterns from exposure and outcome on timescales where the distributed lag is still estimating, bias can occur even when using a distributed lag model.

To illustrate this, one of us (JS) performed a simulation, using the PM_{10} data from Chicago from 1988 to 1993 (to reflect the true serial correlation found in such data) and assuming a true distributed lag between exposure and the log relative risk of death (Figure 3). A Poisson count was simulated for the 6 years of data, 200 times. Figure 4 shows the estimated log relative risk versus degree of freedom per year used to model season (smoothing splines). The two horizontal lines show the true cumulative effect and true effect of lag 0 exposure. The estimated effect at one day falls as progressively more filtering removes the ability of that day's exposure to capture some of the cumulative effect over many days. It appears to asymptote out at 7 *df*/year. This is similar to the reduction in effect size reported by Dominici and colleagues (2002) with an
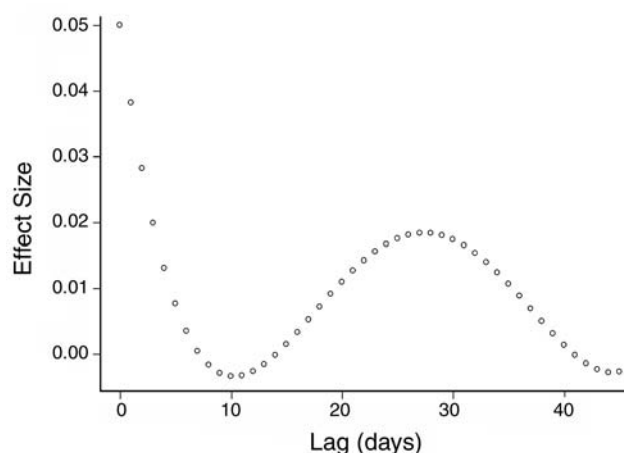
Figure 3. **Chicago (1988–1993).** Postulated distributed lag for PM_{10}.
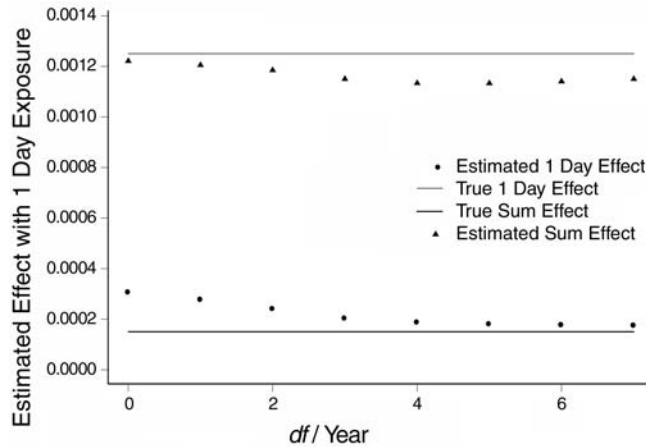
Figure 4. Seasonal control when air pollution effect is spread over multiple days.

increasing degree of freedom, which also asymptoted out at about 7 df/year. Hence, that reduction could be explained by reduced confounding by unknown factors or, equally well, by reduced ability to capture the effect of pollution over multiple days. Interestingly, even when an unconstrained distributed lag is fit and summed to produce an overall effect, there is still a reduction in the estimate. That is, there is bias if a true distributed lag exists.

While one cannot determine from the data alone whether we are reducing true PM effect estimates or removing confounding, it is important to conceptualize what could result in subseasonal patterns correlated with both exposure and outcome. One hypothesized additional confounder that can produce shorter-term excursions in mortality is a respiratory epidemic. Nevertheless, while respiratory epidemics occur in the winter, the timing of when they occur has more to do with how they spread across the world. It seems unlikely that their actual occurrence within a winter is correlated with air pollution. This is a testable hypothesis. Braga and coworkers (2001) examined hospital admissions for pneumonia in multiple cities and used excursions of five-day moving averages of admissions to define epidemic periods. For each day in each epidemic period, they fit a separate 6-degree polynomial to model the impact of the epidemic on mortality. That is, if an epidemic lasted for 17 days, there would be a variable that was zero until the beginning of the epidemic, 1 to 17 for the days of the epidemic, and 17 for all subsequent days. The effect of the epidemic during this period was modeled by a 6-degree polynomial. Separate polynomials were fit for each epidemic similarly. The

epidemics varied in number and intensity from year to year and in intensity from epidemic to epidemic.

After control for all of the epidemic periods by separate 6-degree polynomials, the estimated effect of PM_{10} was unchanged. This lack of change suggests that the most common reason for wanting to fit short-term excursions of mortality is not necessary. Similarly, we previously repeated multiple analyses excluding all days above the 97th percentile of temperature and found no change in PM_{10} coefficients. Therefore, heat waves, which may not be captured by the weather parameters, also seem unlikely as confounders.

On the other hand, fitting more than 3.5 df/year allows one to control for many excursions that are clearly not seasonal and may result in lower estimates of the cumulative effect of environmental variables. This is illustrated in several figures here. Figure 5 shows the results of fitting a 7 df/year spline model to explain seasonality for daily
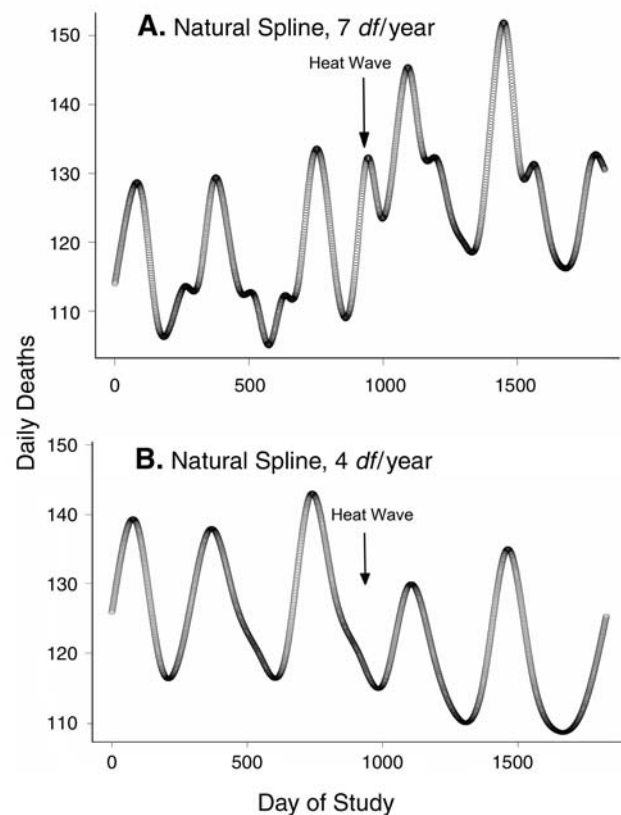
Figure 5. Chicago (1993–1997). Predicted daily deaths over time. *A*. Seasonal model fit using natural spline with 7 df/year. The plot shows a peak of mortality in the summer of 1995, in the middle of the usual summer trough. *B*. Seasonal model fit using natural spline with 4 df/year. The peak of mortality in the summer of 1995 (shown in plot *A*) is explained here by temperature.

deaths in Chicago during the years 1993 to 1997. In the summer of 1995, a serious (and widely studied) heat wave in Chicago was responsible for an estimated 400 early deaths over a 6-day period. Figure 5*A* shows the predicted daily deaths over time from the seasonal model fit to the data. One can see that the 7 *df*/year seasonal model is fitting a peak of mortality in the summer of that year, in the middle of the usual summer trough. That is, 7 *df*/year is sufficient to start picking up not merely season but also short-term (less than a week) changes in other environmental variables. This is despite including a 6-*df* spline for the three weeks around the heat wave itself to capture the effects of the heat wave.

One may argue that the problem is that the heat wave has not been correctly modeled and that the correct attribution would occur if the correct terms and interactions were used. However, the true model is never available. A standard method of attributing deaths to heat waves is to count the excess over the seasonally expected deaths during the heat wave. Clearly, this peak in the summer months attributed to season will reduce the number of excess deaths that can be attributed to the coincident heat wave by that method. The estimated effect of air pollution episodes, such as the London episode of 1952, was similarly computed. Also, if we were somehow able to obtain the correct model for air pollution episodes, weather episodes, and respiratory epidemic episodes, how would we expect the smooth function of time to look? We would expect a basic seasonal pattern, varying in amplitude and phase from year to year, with some shape fluctuations, but not small blips. This is precisely what we get if we use 3 to 4 *df*/year. Because heat waves and epidemics do not appear to be confounders of air pollution, this seems like the appropriate goal for seasonal control. If we want to obtain a good estimate of the effects of short-term fluctuations in environmental variables (weather or pollution), we need to use fewer degrees of freedom for season. This is illustrated in Figure 5*B*, showing the same Chicago data, now fit with 4 *df*/year. In that model, the heat wave is left to be explained by temperature or a heat wave term or by counting excess deaths during the period. That is, imperfect methods of modeling air pollution and weather do better when we do not overattribute short-term fluctuations to season.

Figure 6A shows the predicted daily deaths versus time from Houston for the years 1986 to 1993 using a 7-*df* spline. This model also included splines of temperature, humidity, and indicator variables for day of the week. In

addition to picking up a heat wave in one summer, we see also a short-term blip in the spring of 1986. This short-term excursion of mortality is not seasonality. Whether it is due to weather or to air pollution is not clear. If we want to learn whether short-term excursions in air pollution may be associated with short-term (but not instantaneous) excursions in daily deaths, however, we should not attribute such events to season. Figure 6*B* shows the predicted daily deaths in Houston using a 4-*df* spline, which eliminates the heat wave and the spring excursion from being attributed to season. Finally, Figure 7 shows the partial autocorrelation function for Houston for the original series of daily deaths, for the residuals of the model using 4 *df*/year, and for the residuals of the model using 7 *df*/year. One can see that the daily deaths have positive autocorrelation (Figure 7*A*), that control for season reduced to white noise when 4 *df*/year are used (Figure 7*B*), and that the use
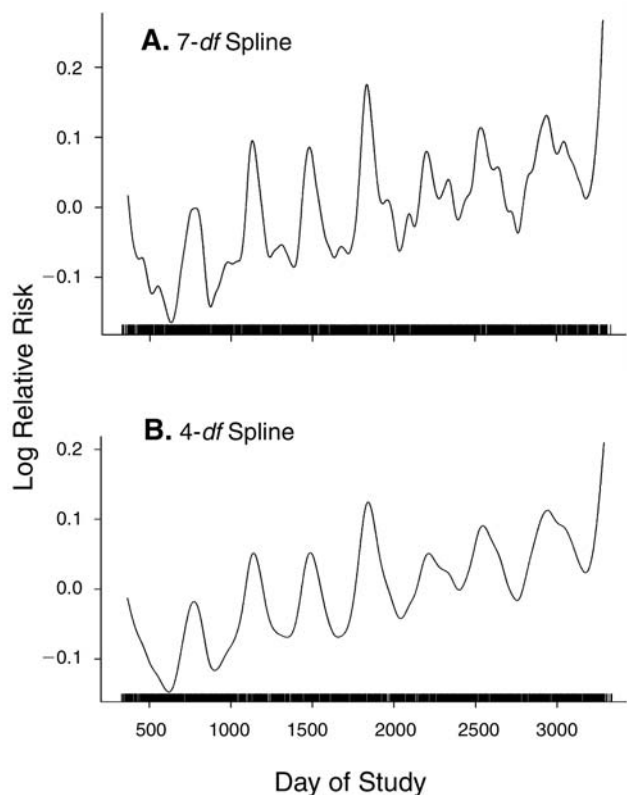
Figure 6. Houston (1986–1993). Predicted daily deaths versus time using 7-*df* (*A*) and 4-*df* (*B*) splines. The model also included splines of temperature, humidity, and indicator variables for day of the week. The 7-*df* spline shows a heat wave in one summer and a very short-term blip in the spring of 1986. The 4-*df* spline eliminates the heat wave and the spring excursion (shown in Figure 4 plot *A*) from being attributed to season.
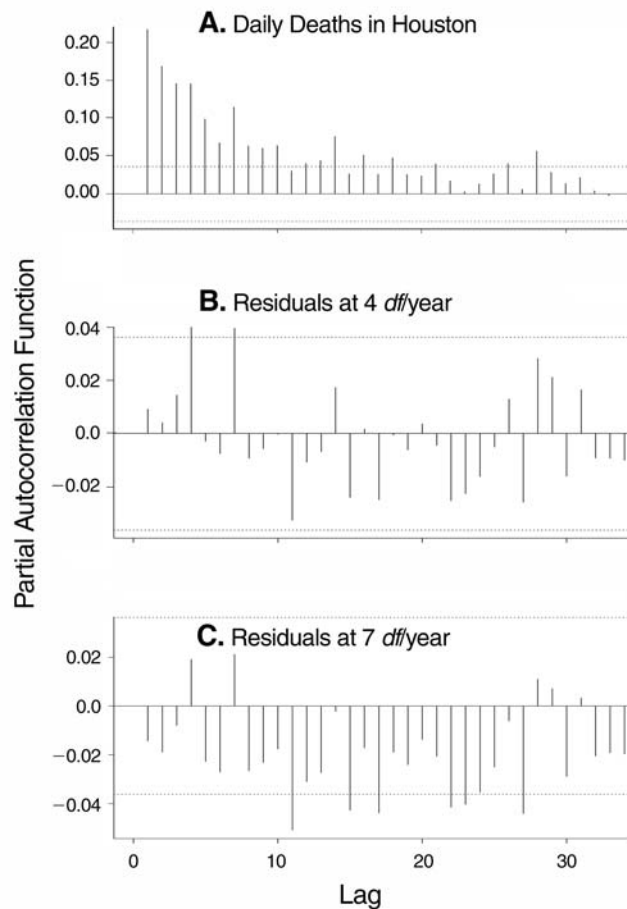
Figure 7. Houston (1986–1993). Partial autocorrelation function. *A*. The original series of daily deaths present a positive autocorrelation; *B,* the model using 4 *df*/year presents residuals reduced to white noise; and *C*, the model using 7 *df*/year shows a negative autocorrelation.

of 7 *df*/year induces negative autocorrelation into the data (Figure 7*C*). This phenomenon is well known in the digital filtering literature, and its occurrence when overfitting time series of health data has been discussed by Diggle (1990). Because air pollution has positive autocorrelation, this phenomenon has the potential to downwardly bias the results for air pollution.

Similarly, we can examine models using different degrees of freedom for temperature to determine reasonable choices. Figure 8 shows the estimated dose-response curves for the effect of today's and yesterday's temperatures on daily deaths in Steubenville Ohio, when using 6 *df* for each curve. These curves are clearly biologically implausible. Indeed, physiologically, we would expect nonlinear but quite smooth curves for temperature. This suggests that 2- or 3-*df* curves are, a priori, more reasonable.
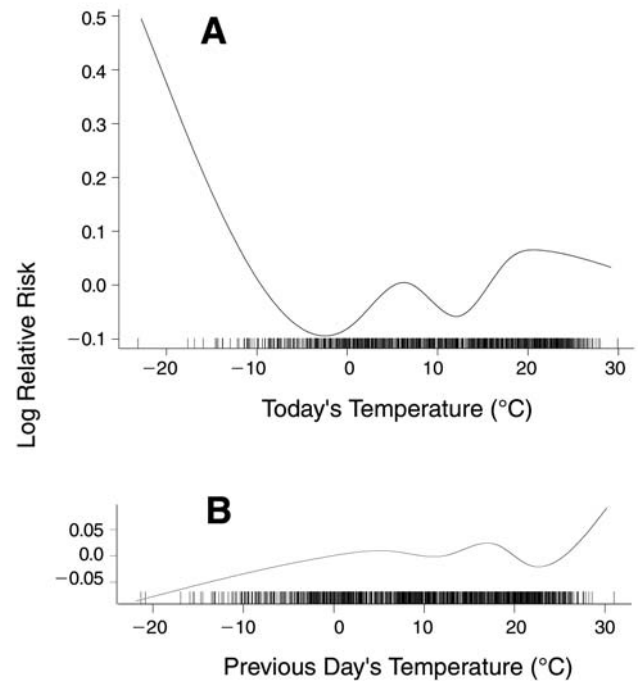
Figure 8. Steubenville, Ohio. The 6 *df* fit for today's temperature (*A*) to daily deaths and for yesterday's temperature (*B*) to daily deaths.

In summary, large degrees of freedom per year are unnecessary to control for season per se, demonstrably fit short-term excursions that may be due to environmental variables, and seem unnecessary to control for respiratory epidemics. On the other hand, given that the effects of air pollution and weather are spread over multiple days, this approach has the potential to downwardly bias the estimated effects of pollution and weather. While controlling for weather is important in air pollution time series, fits with high degrees of freedom produce physiologically implausible curves and may capture the effects of correlated pollution variables. Reason and judgment are needed to define the range of reasonable models to explore.

OTHER ISSUES RELATED TO RESULTS

Choice of Spline

Natural splines are b-splines constrained to be linear at the extreme ranges of the data. Because temperature may have very nonlinear effects at the extremes of its range, we refit all of our models using b-splines instead of n-splines. While estimates varied slightly from city to city, the impact on the metaanalysis of overall effect was trivial. Hence, this does not appear to be a sensitivity of the model.

Standard Errors in Multicity Studies

The estimated standard errors using natural splines were only modestly higher than those with the stricter convergence criteria for GAM. Changes in standard errors for individual cities do not necessarily have much impact on the standard error of the estimated overall effect in a metaanalysis, because that variance includes city-to-city differences as well as the within-city variances. Indeed, if we compare the standard errors for the estimated overall effect from corrected GAMs to natural spline models, we find 0.00010 versus 0.00012 for CVD admissions, 0.00039 versus 0.00038 for COPD admissions, and 0.00028 versus 0.00028 for pneumonia admissions. That is, the results are identical in one instance, higher in another, and lower in the third. A similar result occurred in the reanalysis of NMMAPS mortality study by Dominici and coworkers (elsewhere in this Special Report). The standard error of the estimate is 0.06 with GAM or natural splines. Hence, we believe that in studies such as NMMAPS and APHEA, where there are systematic prospective metaanalyses of many locations, the standard error issue for GAM is unlikely to matter for the combined estimates. If we take into account findings from all studies on the health effects of air pollution, the effect of the standard error problem on overall inference is nil.

The major difference between natural splines and LOESS, then, is not on the standard error of the combined mean estimate across cities (although that is true for individual cities/studies). The major difference is that natural splines fit season and weather differently than does LOESS. This difference in model fit changes the estimate of the mean effect, but not its variance. In NMMAPS, for estimate of effect on mortality across 90 cities, this change was not trivial although all of the major conclusions still held (Dominici et al, in this report). For hospital admissions, the change was modest for CVD and COPD but large for pneumonia.

Sensitivity to Knot Locations and Number

Because the natural splines we used are less flexible, we examined the predicted seasonal and weather patterns for our original models and for the new models in order to identify instances in which they did not seem to similarly fit these patterns.

While parametric approaches like the natural spline have the advantage of no backfitting algorithm or inappropriate estimation of the covariance matrix, the disadvantage of such an approach is that it is less flexible. While by and large the predicted curves versus season and weather
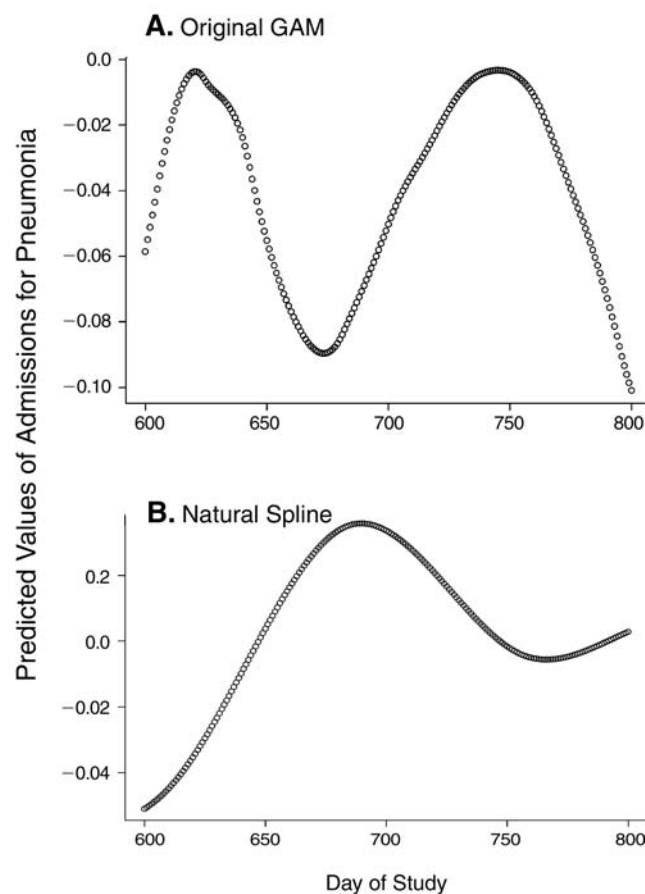
Figure 9. Chicago (September 1989–March 1990). Pneumonia admissions. *A*. Predicted subset plots versus season from the original GAM shows a seasonal pattern with two peaks. *B*. Predicted subset plots versus season from the natural spline model (same degree of freedom as in the GAM) shows only one peak.

were similar, for the smooth function and the splines that was not always the case. For example, Figure 9 shows the predicted plots versus season for pneumonia in Chicago for 200 days in the end of 1989 and beginning of 1990. One plot is from GAM and one uses natural splines with the same degree of freedom. The two models clearly fit different patterns. The plot in Figure 9*A* shows a seasonal pattern with two peaks; the plot in Figure 9*B* shows only one peak. Next Figure 10 shows the deviance residuals from the two models. The plot in Figure 10*A* shows no seasonality left in the data; the plot in Figure 10*B* shows the presence of seasonality in the data. The GAM clearly seems to fit the data better in this case.
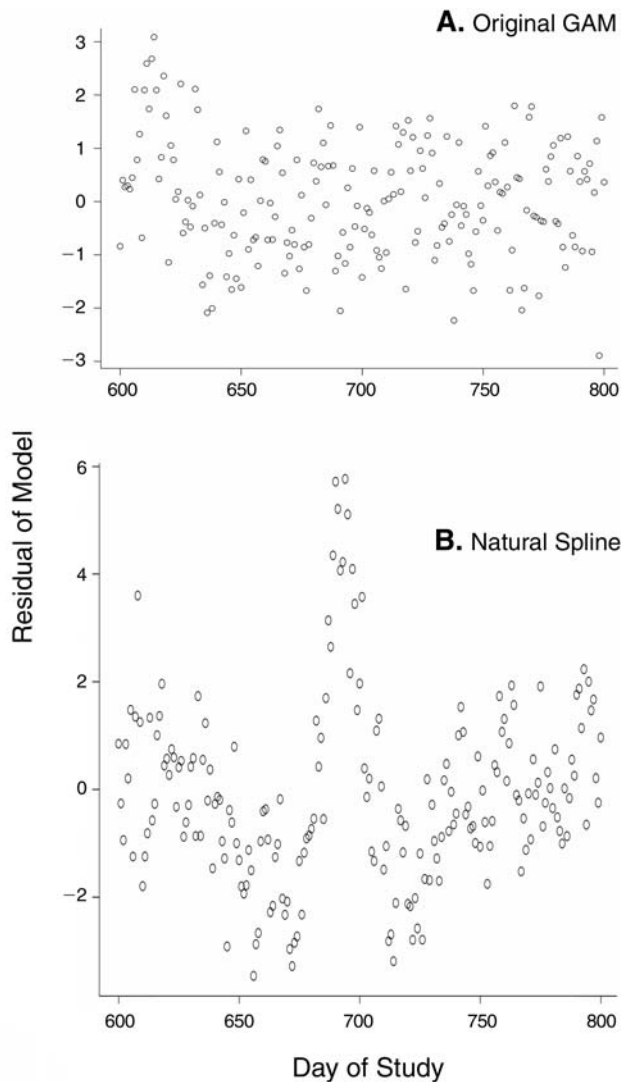
Figure 10. Chicago (September 1989–March 1990). Pneumonia admissions. *A*. Residuals from a subset of the original GAM shows no seasonality left in the data. *B*. Residuals from the natural spline model with the same degree of freedom as in *A*. This plot shows the presence of seasonality in the data.

Another example, given in Figure 1, shows the plots of dose response for temperature in Detroit and pneumonia using 3 and 4 *df*. Increasing the degree of freedom from 3 to

4 completely changed the pattern for temperature. This change in pattern may be due to a different location of the knots over the temperature range.

Effect Modification

Social Factors   Table 6 shows the change from baseline PM_{10} effect (as percentage change in admissions per $10 \ \mu g/m^3$ increase in concentration) associated with a 5-point increase in each measure for models with natural spline and with penalized spline.

For COPD and pneumonia, none of the analyzed factors were significant modifiers of the PM_{10} effect estimates in the 14 cities. For CVD, we found that the percentage of the population living in poverty, the percentage of the population that was nonwhite, and the percentage that was unemployed were modifiers of hospitalizations for CVD. The finding holds for both types of models, even though the results are stronger with natural spline models. The effect of PM_{10} was greater in communities with higher levels of these indicators. This finding differs from our previous results, where they were not significant modifiers, and suggests that social deprivation may convey susceptibility to the effects of particles. This effect could be exerted through increased exposure (eg, because of less air conditioning), increased prevalence of predisposing diseases (eg, diabetes mellitus), or other factors.

Hospitalization Rates   In the metaregression with hospitalization rates, we found that the coefficients for hospital admission rates for CVD, COPD, or pneumonia were not associated with modification of the PM_{10} effect estimates.

Weather   The plot of effect estimates for the distributed lag PM_{10} versus the correlation of PM_{10} with temperature and relative humidity showed similar effects sizes across a broad range of correlations (Figure 11). These results differ little from those previously shown (Samet et al 2000, Figure 32, p 40). In the metaregression we found that the coefficients for temperature and relative humidity were not significant for any of the three outcomes.

Table 6. Effect Modification by Proportion of Population College Educated, Unemployed, Living in Poverty or Nonwhite[a]

	College Educated		Unemployed		Living in Poverty		Nonwhite	
	% Change	95% CI	% Change	95% CI	% Change	95% CI	% Change	95% CI
Natural Spline Model								
CVD	−0.17	−0.42, 0.07	0.78	0.42, 1.14	0.46	0.12, 0.79	0.10	0.01, 0.19
COPD	0.60	−0.47, 1.67	−0.14	−2.60, 2.38	−0.96	−2.65, 0.76	−0.22	−0.65, 0.20
Pneumonia	−0.03	−0.52, 0.47	0.37	−0.70, 1.45	0.46	−0.32, 1.26	0.16	−0.03, 0.34
Penalized Spline Model								
CVD	−0.05	−0.28, 0.17	0.45	0.01, 0.88	0.26	−0.08, 0.60	0.08	0.00, 0.16
COPD	0.87	0.00, 1.76	0.13	−2.23, 2.55	−1.48	−3.14, 0.21	−0.17	−0.62, 0.29
Pneumonia	0.17	−0.31, 0.64	0.24	−0.89, 1.39	0.19	−0.67, 1.06	0.08	−0.13, 0.29

[a] Results are shown for a 10 µg/m^3 increase in PM_{10} and a 5% point increase in effect modifiers in 14 cities.
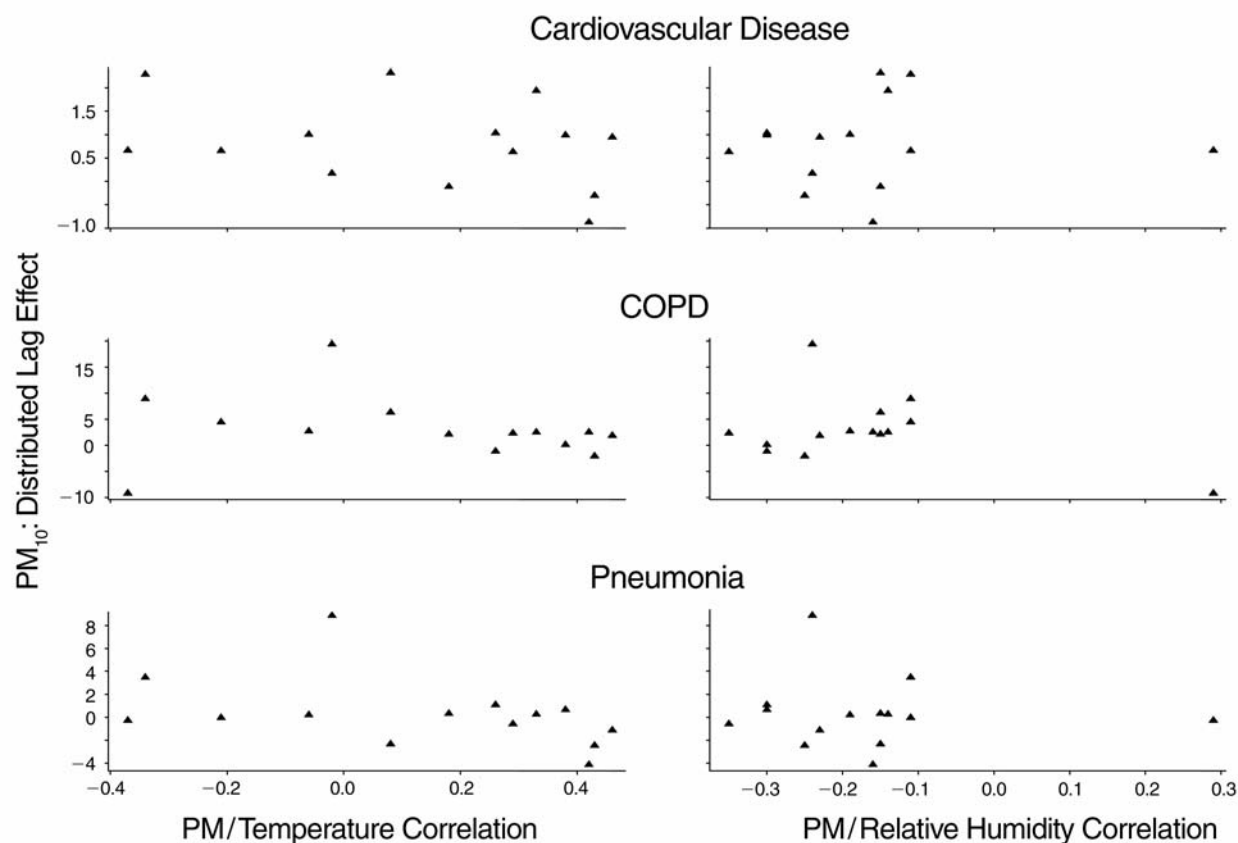
Figure 11. Effect size estimates for the distributed lag of PM_{10} in each city versus the correlation of PM_{10} with temperature and relative humidity for CVD, COPD, and pneumonia.

Copollutants For CVD, COPD, and pneumonia admissions, the effects of PM_{10} in each city have been plotted against the regression coefficients relating SO_2 and O_3 to PM_{10} in each city (Figure 12). These plots show, as seen before (Samet et al 2000, Figure 33, p 41), little evidence of PM_{10} effect confounded by other pollutants except for pneumonia, where control for ozone substantially increased the effect size.

These results have been confirmed by the metaregression estimates shown in Figure 13 for natural spline models and for penalized spline models. Here the baseline estimate is the result of the distributed lag metaanalysis. Plotted above each pollutant is the estimated intercept in the metaregression of the PM_{10} coefficients against the slopes between that copollutant and PM_{10}. The use of metaregression in a second stage to control for confounding by copollutants has been shown to be more resistant to measurement error than traditional two-pollutant models (Schwartz and Coull 2003).
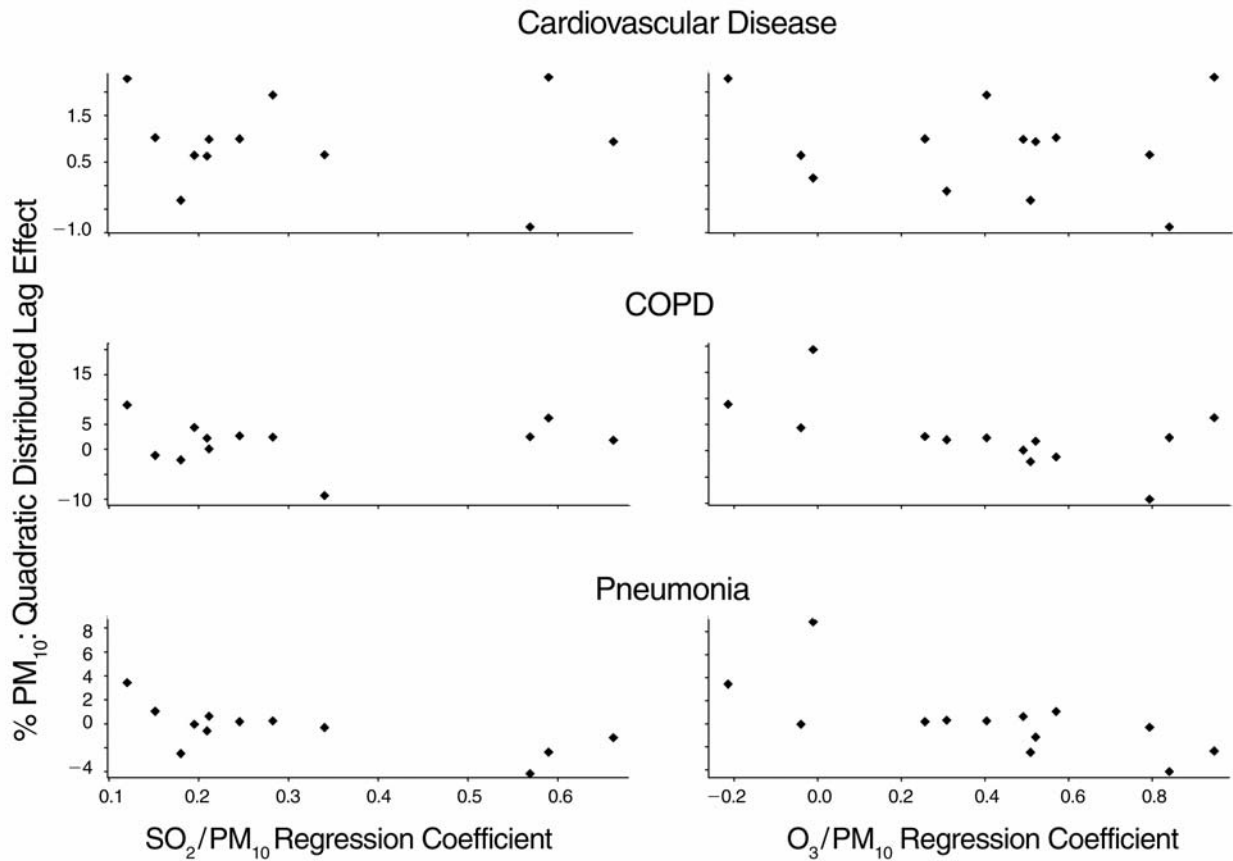
Figure 12. Effect size estimates for quadratic distributed lag of PM_{10} in each city versus the regression coefficients in relation to SO_2 and O_3 to PM_{10} for CVD, COPD, and pneumonia.
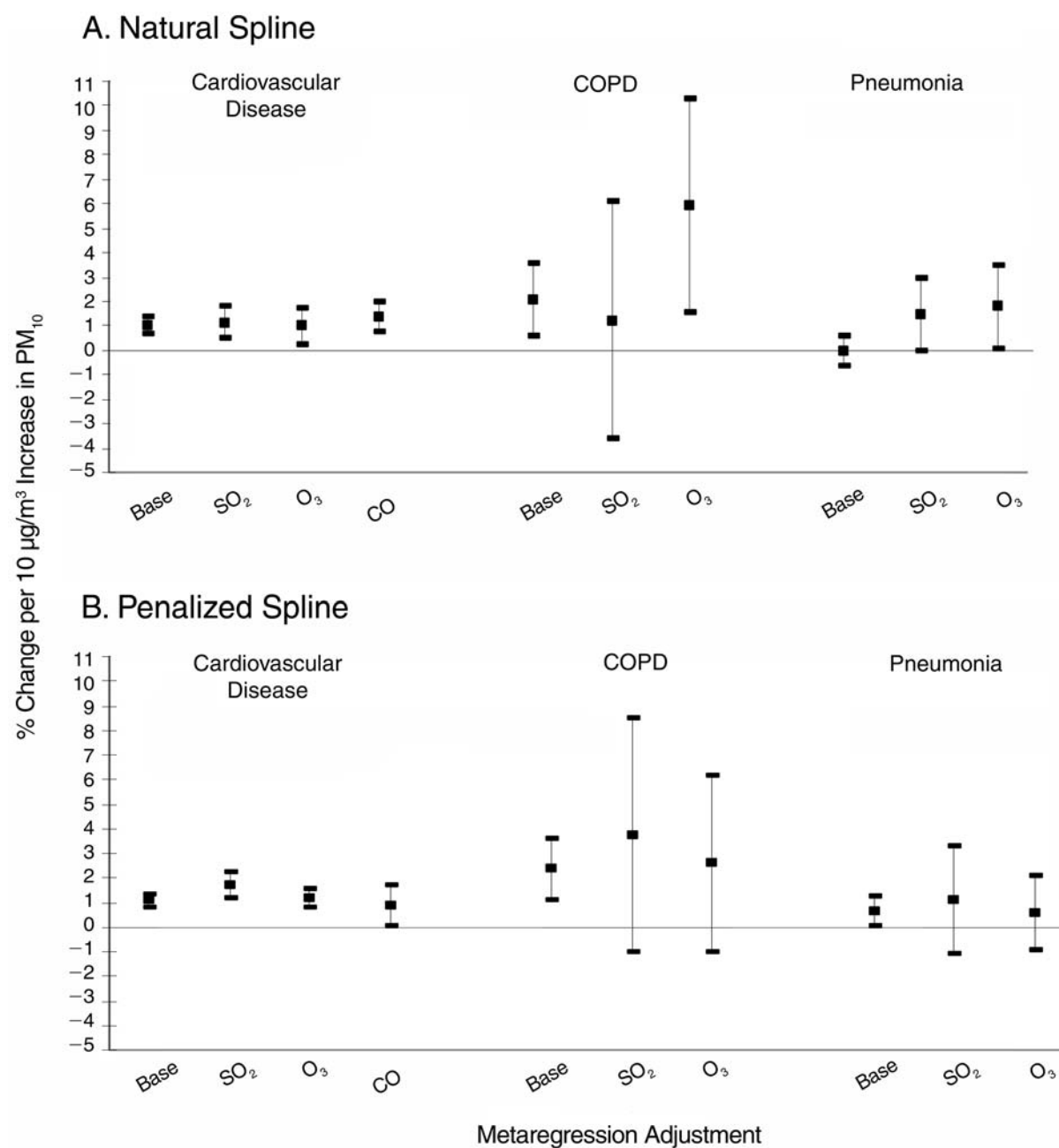
Figure 13. Metaregression adjustment for copollutants. *A.* Natural spline models of combined estimated effects of PM_{10} on cardiovascular disease, COPD and pneumonia without (base) or with adjustment for individual gaseous pollutants. *B.* Penalized spline models of combined estimated effects of PM_{10} on CVD, COPD and pneumonia without (base) or with adjustment for individual gaseous pollutants.

CONCLUSIONS

The major conclusions we believe are warranted are:

1. The S-Plus *gam* function, as written in versions in use through July 2002, has default choices of convergence criteria and iterations that are inappropriate for epidemiologic studies of air pollution.

2. The S-Plus *gam* function, as currently written, does not correctly estimate the standard errors for individual cities although a fix for this will soon be available.

3. Metaanalyses of results from multiple cities (that is, estimates of the mean effect across cities using the original *gam* function), do not appear to have underestimated the standard errors of the results.

4. Several alternative approaches are available to solve problems with the current S-Plus *gam* function. These include: 1) a partial solution using stricter convergence criteria, which in the context of multicity studies is unlikely to compromise the estimated standard errors of overall results; 2) use of an alternative penalized spline GAM, which will avoid the standard error problem, and 3) use of parametric models such as natural spline models or case-crossover analyses. We have reported at least some results using each approach, and all of the data have been reanalyzed with the *gam* function plus tighter convergence criteria or natural splines.

5. With the exception of the pneumonia results, all of the general conclusions previously reported hold true either with stricter convergence criteria in GAM or with regression spline or case-crossover models. In most cases, only minor differences were found among GAM results from with stricter convergence criteria, penalized splines, and natural splines. With pneumonia, however, natural splines showed a large difference compared to GAM with stricter convergence criteria or to penalized splines. In general, reports at the US Environmental Protection Agency (EPA) Workshop on GAM-Related Statistical Issues in PM Epidemiology (November 4–6, 2002, Research Triangle Park NC) also showed that the three methods generally agreed, but that when they differed, the natural spline model separated from the other two. This finding suggests that natural spline models are not as robust as GAMs or penalized spline models. The pneumonia results using natural splines appear to be driven by high pollution days, mostly in the summer when little pneumonia occurs.

6. Air pollution coefficients obtained using regression splines and case-crossover analyses are, on average, lower than those obtained using GAM. The extent of this difference varies considerably from study to study, and from outcome to outcome, and is not uniformly large. While effect estimates have been reduced, the overall effect estimate averaged over multiple studies is usually not dramatically different.

7. Further research is needed on the best ways to control for season and weather; however, conclusions on the effect of exposure to air pollutants may be drawn using the current results from reanalyses of the major multicity studies and from the studies, including multicity studies like APHEA1, which did not use GAMs.

ACKNOWLEDGMENTS

REFERENCES

Bateson TF, Schwartz J. 1999. Control for seasonal variation and time trends in case-crossover studies of acute effects of environmental exposures. Epidemiology 10:539–544.

Bateson TF, Schwartz J. 2001. Selection bias and confounding in case-crossover analyses of environmental time-series data. Epidemiology 12:654–661.

Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA. 1998. Meta-analysis of multiple outcomes by regression with random-effects. Stat Med 17:2537–2550.

Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. 1995. A random-effects regression model for meta-analysis. Stat Med 14:395–411.

Braga ALF, Zanobetti A, Schwartz J. 2001. The lag structure between particulate air pollution and respiratory and cardiovascular deaths in 10 US cities. J Occup Environ Med 43:927–933.

Brumback BA, Ruppert D, Wand MP. 1999. Comment to "Variable selection and function estimation in additive nonparametric regression using data-based prior." J Am Stat Assoc 94:794–797.

Cleveland WS, Devlin SJ. 1988. Robust locally-weighted regression and smoothing scatter plots. J Am Stat Assoc 74:829–836.

Coull BA, Schwartz J, Wand MP. 2001. Respiratory health and air pollution: Additive mixed model analysis. Biostatistics 2:337–349.

Daniels MJ, Dominici F, Samet JM, Zepir SL. 2000. Estimating particulate matter-mortality dose-response curves and threshold levels: An analysis of daily time series for the 20 largest US cities. Am J Epidemiol 152:397-406

Diggle PJ. 1990. Time Series. Oxford University Press, New York NY.

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On the use of Generalized Additive Models in time-series studies of air pollution and health. Am J Epidemiol 156:193–203.

Eilers PHC, Marx BD. 1996. Flexible smoothing with B-splines and penalties (with discussion). Stat Sci 11:89–121.

Goldberg MS, Burnett RT, Brook J, Bailar JC III, Valois MF, Vincent R. 2001. Associations between daily cause-specific mortality and concentrations of ground-level ozone in Montreal, Quebec. Am J Epidemiol 154:817–826.

Hastie TJ, Tibshirani RJ. 1990. Generalized Additive Models. Chapman and Hall, London, England.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiology 12:521–531.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak D, Sunyer J, Bacharova L, Schouten JP, Pontak A, Anderson HR. 1997. Short term effects of ambient sulphur dioxide and particulate mater on mortality in 12 European cities: Results from time series data from the APHEA project. Br Med J 314:1658–1663.

Lee JT, Schwartz J. 1999. Reanalysis of the effects of air pollution on daily mortality in South Korea: A case-crossover design. Environ Health Perspect 107:633–636.

Levy D, Lumley T, Sheppard L, Kaufman J, Checkoway H. 2001. Referent selection in case-crossover analyses of acute health effects of air pollution. Epidemiology 12:186–192.

Levy D, Sheppard L, Checkoway H, Kaufman J, Lumley T, Koenig J, Siscovick D. 2001. A case-crossover analysis of particulate matter air pollution and out of hospital primary cardiac arrest. Epidemiology 12:193–199.

Maclure M. 1991. The case-crossover design: A method for studying transient effects on the risk of acute events. Am J Epidemiol 133:144–153.

Marcus A, Schwartz J. 1987. Dose-response curves for erythrocyte protoporphyrin vs blood lead: Effect of iron status. Environ Res 44:221–227.

McCullagh P, Nelder JA. 1989. Generalized Linear Models (2nd ed). Chapman and Hall, London, England.

Mittleman MA, Maclure M, Tofler GH, Sherwood JB, Goldberg RJ, Muller JE. 1993. Triggering of acute myocardial infarction by heavy physical exertion: Protection by regular exertion. N Engl J Med 329:1677–1683.

Navidi W. 1998. Bidirectional case-crossover designs for exposures with time trends. Biometrics 54:596–605.

Navidi W, Weinhandl E. 2002. Risk set sampling for case-crossover designs. Epidemiology 13:100–105.

Neas LM, Schwartz J, Dockery DW. 1999. A case-crossover analysis of air pollution and mortality in Philadelphia. Environ Health Perspect 107:629–631.

Peters A, Dockery DW, Muller JE, Mittleman MA. 2001. Increased particulate air pollution and the triggering of myocardial infarction. Circulation 103:2810–2815.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Ruppert D, Wand MP, Carroll RJ. 2003. Semiparametric Regression. Cambridge University Press, New York NY.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwartz J, Zanobetti A. 2000. The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Schwartz J. 1993. Air pollution and daily mortality in Birmingham, Alabama. Am J Epidemiol 137:1136–1147.

Schwarz J. 1994. Nonparametric smoothing in the analysis of air pollution and respiratory illness. Can J Stat 22:471–487.

Schwartz J. 2000. Assessing confounding, effect modification, and thresholds in the association between ambient particles and daily deaths. Environ Health Perspect 108:563–568.

Schwartz J, Coull B. 2003. Control for confounding in the presence of measurement error in hierarchical models. Biostatistics. In press.

Schwartz J, Dockery DW. 1992a. Particulate air pollution and daily mortality in Steubenville, Ohio. Am J Epidemiol 135:12–20.

Schwartz J, Dockery DW. 1992b. Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 145:600–604.

Sunyer J, Schwartz J, Tobias A, Macfalane D, Garcia J, Anto JM. 2000. Patients with chronic obstructive pulmonary disease are at increased risk of death associated with urban particulate air pollution: A case-crossover analysis. Am J Epidemiol 151:50–56.

Wood SN. 2000. Modeling and smoothing parameter estimation with multiple quadratic penalties. J R Stat Soc Ser B 62:413–428.

Zelikoff JT, Nadziejko C, Fang T, Gordon C, Premdass C, Cohen MD. 1999. Short term, low-dose inhalation of ambient particulate matter exacerbates ongoing pneumococcal infections in *Streptococcus pneumoniae*-infected rats. In: Proceedings of the Third Colloquium on Particulate Air Pollution and Human Health (Phalen RF, Bell YM, eds), pp 8-94–8-101. Air Pollution Health Effects Laboratory, University of California, Irvine CA.

APPENDIX A. City-Specific Evaluations

HOSPITAL ADMISSIONS AND PM$_{10}$ LEVELS

The details of the original models for each outcome and each city are given in the original NMMAPS report in Appendix D, Tables D.2 through D.4 (Samet et al 2000, pp 65–66). Tables A.1 through A.3 present the new results for hospital admissions, for the mean of PM$_{10}$ on the day of and day before the admission (lag 0/1) and for the distributed lag models, using GAM with the stricter convergence criteria. The metaanalysis results are for the random effects model.

Tables A.4 through A.6 show the city-specific models with the same degree of freedom used with natural spline; these correspond to Tables D.2 through D.4 in the original NMMAPS report (Samet et al 2000, pp 65–66). Tables A.7 through A.9 show the results of the models using natural spline for PM$_{10}$ lag 0, lag 1, two-day mean, and quadratic and unconstrained distributed lag models. Table A.10 presents results for the two-day mean for PM$_{10}$ less than 50 µg/m^3. These tables correspond to Tables D.5 through D.8 in the Appendix D of the original NMMAPS report (Samet et al 2000, pp 67–68).

Tables A.11 through A.13 show the city-specific results of models using penalized spline for the two-day mean of PM$_{10}$ and for quadratic and unconstrained distributed lag models. The models were fit using the same degrees of freedom as in the original GAM with LOESS.

MORTALITY AND PM$_{10}$ LEVELS

Table A.14 reports the results of reanalyzing the distributed lag mortality data in each city, also either a) changing only the convergence criteria or b) using natural spline models with the same degree of freedom as before (Samet et al 2000, Table B.3, p 58). Appendix B presents city-specific and mean results from the case-crossover approach (Table B.3, this report).

Table A.1. City-Specific and Combined Analyses of PM_{10} Associations with CVD Admissions[a]

City	PM$_{10}$ Lag 0/1			Quadratic Distributed Lag			Unrestricted Distributed Lag		
	β	SE	t	β	SE	t	β	SE	t
Birmingham	0.00042	0.00032	1.3	0.00095	0.00047	2.0	0.00102	0.00047	2.2
Boulder	0.00175	0.00140	1.2	−0.00067	0.00210	−0.3	−0.00094	0.00211	−0.4
Canton	0.00067	0.00068	1.0	−0.00137	0.00113	−1.2	−0.00131	0.00113	−1.2
Chicago	0.00102	0.00016	6.6	0.00108	0.00023	4.8	0.00107	0.00023	4.7
Colorado Springs	0.00114	0.00090	1.3	0.00227	0.00138	1.6	0.00220	0.00139	1.6
Detroit	0.00120	0.00016	7.6	0.00160	0.00023	6.8	0.00160	0.00023	6.9
Minneapolis	0.00071	0.00038	1.9	0.00057	0.00059	1.0	0.00062	0.00060	1.0
Nashville	0.00018	0.00060	0.3	−0.00035	0.00103	−0.3	−0.00031	0.00103	−0.3
New Haven	0.00212	0.00042	5.1	0.00204	0.00067	3.0	0.00205	0.00067	3.0
Pittsburgh	0.00113	0.00017	6.7	0.00113	0.00025	4.5	0.00114	0.00025	4.6
Provo/Orem	0.00033	0.00057	0.6	0.00072	0.00087	0.8	0.00075	0.00087	0.9
Seattle	0.00104	0.00026	4.0	0.00130	0.00036	3.6	0.00135	0.00036	3.7
Spokane	0.00060	0.00033	1.8	0.00073	0.00038	1.9	0.00077	0.00038	2.0
Youngstown	0.00100	0.00062	1.6	0.00019	0.00103	0.2	0.00019	0.00104	0.2
Maximum likelihood	0.00099	0.00010	9.8	0.00109	0.00014	7.6	0.00111	0.00014	7.9

[a] GAM with LOESS and stricter convergence criteria.

Table A.2. City-Specific and Combined Results of Analysis of PM_{10} Associations with COPD Admissions[a]

City	PM$_{10}$ Lag 0/1			Quadratic Distributed Lag			Unrestricted Distributed Lag		
	β	SE	t	β	SE	t	β	SE	t
Birmingham	−0.00148	0.00108	−1.4	−0.00131	0.00161	−0.8	−0.00118	0.00161	−0.7
Boulder	0.01117	0.00372	3.0	0.01867	0.00561	3.3	0.01788	0.00564	3.2
Canton	0.00148	0.00185	0.8	0.00284	0.00304	0.9	0.00279	0.00305	0.9
Chicago	0.00145	0.00051	2.8	0.00086	0.00083	1.0	0.00084	0.00083	1.0
Colorado Springs	0.00076	0.00250	0.3	0.00713	0.00347	2.1	0.00771	0.00348	2.2
Detroit	0.00207	0.00049	4.2	0.00300	0.00077	3.9	0.00297	0.00077	3.8
Minneapolis	0.00341	0.00118	2.9	0.00352	0.00184	1.9	0.00334	0.00185	1.8
Nashville	0.00214	0.00191	1.1	0.00240	0.00308	0.8	0.00198	0.00309	0.6
New Haven	0.00374	0.00177	2.1	0.00679	0.00276	2.5	0.00675	0.00277	2.4
Pittsburgh	0.00233	0.00047	5.0	0.00262	0.00073	3.6	0.00245	0.00074	3.3
Provo/Orem	−0.00296	0.00260	−1.1	−0.00732	0.00396	−1.8	−0.00713	0.00397	−1.8
Seattle	0.00099	0.00083	1.2	0.00443	0.00113	3.9	0.00459	0.00113	4.1
Spokane	0.00198	0.00097	2.0	0.00259	0.00102	2.5	0.00271	0.00103	2.6
Youngstown	0.00143	0.00181	0.8	−0.00195	0.00316	−0.6	−0.00162	0.00318	−0.5
Maximum likelihood	0.00170	0.00039	4.4	0.00250	0.00067	3.7	0.00250	0.00066	3.8

[a] GAM with LOESS and stricter convergence criteria.

Table A.3. City-Specific and Combined Results of Analysis of PM_{10} Associations with Pneumonia Admissions[a]

City	PM_{10} Lag 0/1			Quadratic Distributed Lag			Unrestricted Distributed Lag		
	β	SE	t	β	SE	t	β	SE	t
Birmingham	0.00016	0.00061	0.3	0.00056	0.00083	0.7	0.00061	0.00083	0.7
Boulder	0.00354	0.00259	1.4	0.00535	0.00363	1.5	0.00507	0.00364	1.4
Canton	0.00115	0.00140	0.8	−0.00131	0.00228	−0.6	−0.00127	0.00229	−0.6
Chicago	0.00201	0.00030	6.6	0.00206	0.00045	4.6	0.00274	0.00046	6.0
Colorado Springs	0.00439	0.00152	2.9	0.00814	0.00225	3.6	0.00792	0.00226	3.5
Detroit	0.00266	0.00033	8.0	0.00123	0.00050	2.5	0.00113	0.00050	2.3
Minneapolis	0.00291	0.00071	4.1	0.00179	0.00108	1.6	0.00190	0.00109	1.7
Nashville	−0.00126	0.00127	−1.0	0.00065	0.00189	0.3	0.00071	0.00189	0.4
New Haven	0.00287	0.00082	3.5	−0.00141	0.00135	−1.0	−0.00067	0.00138	−0.5
Pittsburgh	0.00170	0.00035	4.8	0.00126	0.00054	2.3	0.00126	0.00054	2.3
Provo/Orem	0.00035	0.00100	0.3	0.00279	0.00141	2.0	0.00294	0.00141	2.1
Seattle	0.00148	0.00051	2.9	0.00166	0.00069	2.4	0.00150	0.00070	2.1
Spokane	0.00081	0.00063	1.3	0.00156	0.00068	2.3	0.00171	0.00069	2.5
Youngstown	0.00156	0.00137	1.1	0.00077	0.00220	0.3	0.00095	0.00222	0.4
Maximum likelihood	0.00169	0.00028	6.1	0.00146	0.00031	4.7	0.00160	0.00034	4.8

[a] GAM with LOESS and stricter convergence criteria.

Table A.4. Base Model Characteristics for CVD Admissions[a]

City	Season	Temperature	Relative Humidity	Temperature Lag 1	Barometric Pressure	Day of Week	Autoregressive Terms
Birmingham	20	3	3	3	4	6	none
Boulder	6	2	2	3	2	6	none
Canton	11	3	2	3	4	6	none
Chicago	19	3	2	2	4	6	ar1,ar2,ar3
Colorado Springs	10	3	3	3	3	6	none
Detroit	17	3	4	3	4	6	ar1,ar2,ar3
Minneapolis	21	3	4	3	3	6	none
Nashville	4	2	3	2	2	6	none
New Haven	13	2	3	3	4	6	none
Pittsburgh	26	3	2	3	4	6	ar1
Provo/Orem	8	3	3	4	3	6	none
Seattle	28	3	2	3	4	6	none
Spokane	18	3	2	3	3	6	none
Youngstown	12	3	3	3	2	6	none

[a] Units are degrees of freedom.

Table A.5. Base Model Characteristics for COPD Admissions[a]

City	Season	Temperature	Relative Humidity	Temperature Lag 1	Barometric Pressure	Day of Week	Autoregressive Terms
Birmingham	18	2	2	2	3	6	none
Boulder	10	3	2	2	2	6	none
Canton	10	3	2	3	3	6	none
Chicago	24	3	2	2	3	6	ar1,ar2,ar3
Colorado Springs	14	3	3	2	2	6	none
Detroit	20	3	3	3	3	6	ar1,ar2,ar3
Minneapolis	24	3	3	3	3	6	none
Nashville	14	3	3	3	4	6	none
New Haven	6	3	2	3	4	6	none
Pittsburgh	32	3	2	3	4	6	none
Provo/Orem	8	2	2	2	2	6	none
Seattle	21	3	4	3	5	6	none
Spokane	17	3	2	3	3	6	none
Youngstown	7	2	5	2	3	6	none

[a] Units are degrees of freedom.

Table A.6. Base Model Characteristics for Pneumonia Admissions[a]

City	Season	Temperature	Relative Humidity	Temperature Lag 1	Barometric Pressure	Day of Week	Autoregressive Terms
Birmingham	22	5	3	5	3	6	ar1,ar2,ar4,ar5,ar7
Boulder	11	5	3	3	3	3	none
Canton	17	3	2	3	3	6	none
Chicago	35	2	2	2	3	6	ar1,ar2,ar3
Colorado Springs	24	2	4	2	2	3	none
Detroit	22	3	2	4	4	6	ar1,ar2,ar3,ar4
Minneapolis	29	3	3	3	3	6	ar1,ar2,ar3
Nashville	14	3	3	3	2	3	ar1,ar2
New Haven	18	3	3	4	5	6	ar1
Pittsburgh	36	4	2	4	5	6	ar1,ar2,ar3,ar4
Provo/Orem	19	2	3	2	3	6	none
Seattle	26	5	4	5	3	6	ar1,ar2,ar3
Spokane	27	3	2	3	3	6	none
Youngstown	14	4	5	3	5	6	none

[a] Units are degrees of freedom.

Table A.7. City-Specific and Combined Results of Analysis of PM_{10} Associations with CVD Admissions[a]

City	PM_{10} Lag 0			PM_{10} Lag 1			PM_{10} Lag 0/1			Quadratic Distributed Lag			Unrestricted Distributed Lag		
	β	SE	t	β	SE	t	β	SE	t	β	SE	t	β	SE	t
Birmingham	0.00069	0.00030	2.3	0.000044	0.000293	0.2	0.00047	0.00034	1.4	0.00102	0.00050	2.0	0.00108	0.00050	2.2
Boulder	0.00301	0.00141	2.1	0.000009	0.001385	0.0	0.00218	0.00154	1.4	0.00016	0.00238	0.1	−0.00010	0.00239	0.0
Canton	0.00034	0.00081	0.4	0.000849	0.000749	1.1	0.00083	0.00076	1.1	−0.00088	0.00132	−0.7	−0.00081	0.00132	−0.6
Chicago	0.00101	0.00017	6.0	0.000533	0.000167	3.2	0.00087	0.00017	5.0	0.00099	0.00025	4.0	0.00113	0.00025	4.6
Colorado Springs	0.00021	0.00092	0.2	0.001353	0.000857	1.6	0.00117	0.00094	1.2	0.00228	0.00149	1.5	0.00221	0.00149	1.5
Detroit	0.00152	0.00019	7.8	0.001153	0.000183	6.3	0.00146	0.00018	8.3	0.00193	0.00028	6.9	0.00193	0.00028	6.8
Minneapolis	0.00113	0.00038	2.9	0.000195	0.000370	0.5	0.00074	0.00042	1.7	0.00063	0.00065	1.0	0.00065	0.00065	1.0
Nashville	−0.00021	0.00061	−0.3	0.000055	0.000606	0.1	0.00027	0.00063	0.4	−0.00012	0.00106	−0.1	−0.00008	0.00106	−0.1
New Haven	0.00197	0.00047	4.2	0.001075	0.000466	2.3	0.00221	0.00050	4.4	0.00231	0.00084	2.7	0.00229	0.00084	2.7
Pittsburgh	0.00084	0.00017	5.0	0.000631	0.000165	3.8	0.00099	0.00019	5.1	0.00094	0.00029	3.2	0.00095	0.00029	3.3
Provo/Orem	0.00052	0.00064	0.8	0.000006	0.000578	0.0	0.00030	0.00067	0.4	0.00066	0.00097	0.7	0.00069	0.00097	0.7
Seattle	0.00100	0.00028	3.5	0.000205	0.000277	0.7	0.00068	0.00030	2.2	0.00065	0.00044	1.5	0.00070	0.00044	1.6
Spokane	0.00069	0.00034	2.0	0.000629	0.000314	2.0	0.00077	0.00035	2.2	0.00100	0.00040	2.5	0.00104	0.00041	2.6
Youngstown	0.00173	0.00071	2.4	−0.000585	0.000689	−0.8	0.00066	0.00072	0.9	−0.00031	0.00122	−0.3	−0.00038	0.00122	−0.3
Maximum likelihood	0.00101	0.00013	8.0	0.000520	0.000120	4.3	0.00095	0.00012	7.7	0.00105	0.00017	6.2	0.00111	0.00016	6.7

[a] Natural spline model.

Table A.8. City-Specific and Combined Results of Analysis of PM_{10} Associations with COPD Admissions[a]

City	PM_{10} Lag 0			PM_{10} Lag 1			PM_{10} Lag 0/1			Quadratic Distributed Lag			Unrestricted Distributed Lag		
	β	SE	t	β	SE	t	β	SE	t	β	SE	t	β	SE	t
Birmingham	−0.00075	0.00100	−0.7	−0.00144	0.00098	−1.5	−0.00132	0.00114	−1.2	−0.00120	0.00172	−0.7	−0.00109	0.00173	−0.6
Boulder	0.00732	0.00383	1.9	0.00958	0.00363	2.6	0.01145	0.00420	2.7	0.01934	0.00657	2.9	0.01861	0.00659	2.8
Canton	−0.00004	0.00221	0.0	0.00361	0.00199	1.8	0.00115	0.00207	0.6	0.00249	0.00360	0.7	0.00235	0.00361	0.6
Chicago	0.00045	0.00053	0.8	0.00066	0.00052	1.3	0.00073	0.00058	1.3	0.00010	0.00090	0.1	0.00007	0.00090	0.1
Colorado Springs	0.00279	0.00250	1.1	−0.00129	0.00246	−0.5	0.00164	0.00268	0.6	0.00891	0.00376	2.4	0.00950	0.00377	2.5
Detroit	0.00151	0.00064	2.3	0.00187	0.00060	3.1	0.00174	0.00058	3.0	0.00248	0.00098	2.5	0.00248	0.00099	2.5
Minneapolis	0.00176	0.00120	1.5	0.00181	0.00118	1.5	0.00263	0.00134	2.0	0.00224	0.00205	1.1	0.00214	0.00205	1.0
Nashville	0.00154	0.00194	0.8	0.00088	0.00198	0.4	0.00198	0.00201	1.0	0.00212	0.00326	0.6	0.00168	0.00328	0.5
New Haven	0.00442	0.00189	2.3	0.00207	0.00183	1.1	0.00356	0.00198	1.8	0.00629	0.00314	2.0	0.00622	0.00315	2.0
Pittsburgh	0.00124	0.00047	2.6	0.00112	0.00046	2.4	0.00159	0.00055	2.9	0.00182	0.00086	2.1	0.00166	0.00087	1.9
Provo/Orem	−0.00363	0.00287	−1.3	−0.00413	0.00258	−1.6	−0.00473	0.00298	−1.6	−0.00927	0.00434	−2.1	−0.00921	0.00435	−2.1
Seattle	0.00083	0.00088	0.9	0.00022	0.00086	0.3	0.00073	0.00094	0.8	0.00439	0.00129	3.4	0.00456	0.00129	3.5
Spokane	0.00212	0.00100	2.1	0.00168	0.00093	1.8	0.00196	0.00101	1.9	0.00270	0.00105	2.6	0.00282	0.00106	2.7
Youngstown	0.00034	0.00208	0.2	0.00139	0.00185	0.7	0.00177	0.00199	0.9	−0.00208	0.00360	−0.6	−0.00165	0.00361	−0.5
Maximim likelihood	0.00108	0.00031	3.5	0.00098	0.00038	2.6	0.00131	0.00038	3.4	0.00208	0.00074	2.8	0.00209	0.00074	0.0

[a] Natural spline model.

Table A.9. City-Specific and Combined Results of Analysis of PM$_{10}$ Associations with Pneumonia Admissions[a]

City	PM$_{10}$ Lag 0			PM$_{10}$ Lag 1			PM$_{10}$ Lag 0/1			Quadratic Distributed Lag			Unrestricted Distributed Lag		
	β	SE	t	β	SE	t	β	SE	t	β	SE	t	β	SE	t
Birmingham	0.00029	0.00060	0.5	0.00004	0.00057	0.1	0.00000	0.00065	0.0	0.00105	0.00089	1.2	0.00112	0.00089	1.3
Boulder	0.00635	0.00260	2.4	0.00133	0.00261	0.5	0.00512	0.00290	1.8	0.00884	0.00432	2.0	0.00897	0.00433	2.1
Canton	−0.00014	0.00171	−0.1	−0.00101	0.00165	−0.6	−0.00057	0.00165	−0.3	−0.00417	0.00282	−1.5	−0.00410	0.00282	−1.5
Chicago	0.00076	0.00032	2.4	0.00007	0.00031	0.2	0.00082	0.00033	2.4	0.00062	0.00052	1.2	0.00080	0.00052	1.5
Colorado Springs	0.00211	0.00163	1.3	0.00122	0.00156	0.8	0.00249	0.00172	1.4	0.00345	0.00273	1.3	0.00318	0.00273	1.2
Detroit	0.00148	0.00044	3.4	0.00076	0.00042	1.8	0.00184	0.00039	4.7	0.00024	0.00060	0.4	0.00015	0.00060	0.2
Minneapolis	0.00098	0.00077	1.3	0.00085	0.00076	1.1	0.00115	0.00080	1.4	−0.00060	0.00125	−0.5	−0.00058	0.00122	−0.5
Nashville	0.00111	0.00151	0.7	0.00085	0.00159	0.5	−0.00143	0.00135	−1.1	0.00030	0.00201	0.2	0.00041	0.00202	0.2
New Haven	0.00145	0.00100	1.4	−0.00228	0.00100	−2.3	0.00121	0.00100	1.2	−0.00236	0.00160	−1.5	−0.00226	0.00160	−1.4
Pittsburgh	0.00047	0.00036	1.3	−0.00064	0.00035	−1.8	−0.00009	0.00041	−0.2	−0.00116	0.00064	−1.8	−0.00119	0.00064	−1.8
Provo/Orem	−0.00003	0.00115	0.0	−0.00067	0.00107	−0.6	−0.00046	0.00123	−0.4	−0.00030	0.00134	−0.2	−0.00001	0.00135	0.0
Seattle	0.00048	0.00058	0.8	0.00068	0.00056	1.2	0.00062	0.00059	1.0	−0.00004	0.00082	−0.1	−0.00020	0.00083	−0.2
Spokane	0.00044	0.00068	0.6	−0.00073	0.00065	−1.1	−0.00035	0.00070	−0.5	0.00017	0.00076	0.2	0.00033	0.00076	0.4
Youngstown	0.00133	0.00154	0.9	−0.00130	0.00153	−0.8	−0.00052	0.00158	−0.3	−0.00248	0.00263	−0.9	−0.00236	0.00265	−0.9
Maximum likelihood	0.00077	0.00017	4.6	−0.00003	0.00022	−0.1	0.00056	0.00027	2.1	−0.00002	0.00032	−0.1	0.00001	0.00033	0.0

[a] Natural spline model.

Table A.10. Results for Two-Day Mean Exposure to PM$_{10}$ for Values Less Than 50 µg/m^3 and Hospital Admissions for Specific Diagnoses in 14 US Cities[a]

City	CVD			COPD			Pneumonia		
	β	SE	t	β	SE	t	β	SE	t
Birmingham	0.00041	0.00061	0.7	−0.00273	0.00200	−1.4	−0.00221	0.00153	−1.4
Boulder	0.00015	0.00197	0.1	0.00871	0.00536	1.6	0.00550	0.00360	1.5
Canton	0.00104	0.00099	1.1	0.00043	0.00268	0.2	−0.00113	0.00212	−0.5
Chicago	0.00058	0.00035	1.7	0.00156	0.00107	1.5	0.00128	0.00062	2.1
Colorado Springs	0.00208	0.00150	1.4	0.00436	0.00424	1.0	0.00170	0.00291	0.6
Detroit	0.00248	0.00039	6.3	0.00213	0.00131	1.6	0.00249	0.00091	2.7
Minneapolis	0.00105	0.00058	1.8	0.00252	0.00184	1.4	0.00266	0.00116	2.3
Nashville	0.00044	0.00090	0.5	0.00533	0.00281	1.9	−0.00105	0.00199	−0.5
New Haven	0.00345	0.00072	4.8	−0.00043	0.00300	−0.1	0.00173	0.00152	1.1
Pittsburgh	0.00181	0.00037	4.9	0.00214	0.00107	2.0	0.00169	0.00094	1.8
Provo/Orem	0.00190	0.00154	1.2	−0.00555	0.00725	−0.8	0.00121	0.00311	0.4
Seattle	0.00050	0.00056	0.9	0.00405	0.00168	2.4	0.00259	0.00122	2.1
Spokane	0.00072	0.00088	0.8	0.00370	0.00249	1.5	−0.00365	0.00164	−2.2
Youngstown	0.00113	0.00109	1.0	0.00722	0.00303	2.4	0.00101	0.00232	0.4
Maximum likelihood	0.00131	0.00028	4.7	0.00219	0.00060	3.7	0.00105	0.00051	2.1

[a] Natural spline model.

Table A.11. City-Specific and Combined Results of Analysis of PM_{10} Associations with CVD Admissions[a]

City	PM_{10} Lag 0/1			Quadratic Distributed Lag			Unconstrained Distributed Lag		
	β	SE	t	β	SE	t	β	SE	t
Birmingham	0.00062	0.00035	1.8	0.00128	0.00050	2.6	0.00133	0.00050	2.7
Boulder	0.00160	0.00157	1.0	−0.00095	0.00233	−0.4	−0.00116	0.00233	−0.5
Canton	0.00108	0.00077	1.4	−0.00116	0.00129	−0.9	−0.00107	0.00129	−0.8
Chicago	0.00103	0.00015	6.8	0.00109	0.00024	4.6	0.00109	0.00024	4.6
Colorado Springs	0.00106	0.00099	1.1	0.00233	0.00145	1.6	0.00226	0.00146	1.6
Detroit	0.00119	0.00016	7.2	0.00161	0.00028	5.8	0.00171	0.00028	6.2
Minneapolis	0.00068	0.00042	1.6	0.00040	0.00062	0.6	0.00044	0.00062	0.7
Nashville	0.00053	0.00066	0.8	−0.00028	0.00105	−0.3	−0.00029	0.00105	−0.3
New Haven	0.00242	0.00048	5.0	0.00233	0.00078	3.0	0.00232	0.00078	3.0
Pittsburgh	0.00099	0.00019	5.3	0.00100	0.00028	3.6	0.00102	0.00028	3.6
Provo/Orem	0.00040	0.00073	0.6	0.00121	0.00077	1.6	0.00128	0.00078	1.6
Seattle	0.00087	0.00031	2.8	0.00100	0.00041	2.4	0.00103	0.00042	2.5
Spokane	0.00047	0.00033	1.5	0.00065	0.00039	1.7	0.00070	0.00040	1.7
Youngstown	0.00113	0.00072	1.6	0.00061	0.00117	0.5	0.00058	0.00118	0.5
Maximum likelihood	0.00099	0.00010	10.1	0.00108	0.00014	7.7	0.00111	0.00015	7.6

[a] Penalized spline model.

Table A.12. City-Specific and Combined Results of Analysis of PM_{10} Associations with COPD Admissions[a]

City	PM_{10} Lag 0/1			Quadratic Distributed Lag			Unconstrained Distributed Lag		
	β	SE	t	β	SE	t	β	SE	t
Birmingham	−0.00112	0.00118	−0.9	−0.00104	0.00168	−0.6	−0.00094	0.00168	−0.6
Boulder	0.01136	0.00405	2.8	0.01801	0.00642	2.8	0.01742	0.00643	2.7
Canton	0.00075	0.00208	0.4	0.00107	0.00347	0.3	0.00094	0.00348	0.3
Chicago	0.00121	0.00057	2.1	0.00109	0.00088	1.2	0.00108	0.00088	1.2
Colorado Springs	0.00142	0.00256	0.6	0.00834	0.00371	2.3	0.00855	0.00372	2.3
Detroit	0.00180	0.00057	3.2	0.00304	0.00095	3.2	0.00411	0.00098	4.2
Minneapolis	0.00293	0.00131	2.2	0.00321	0.00194	1.7	0.00308	0.00194	1.6
Nashville	0.00178	0.00200	0.9	0.00182	0.00321	0.6	0.00140	0.00323	0.4
New Haven	0.00356	0.00190	1.9	0.00619	0.00305	2.0	0.00613	0.00305	2.0
Pittsburgh	0.00191	0.00054	3.5	0.00212	0.00083	2.6	0.00198	0.00083	2.4
Provo/Orem	−0.00356	0.00280	−1.3	−0.00225	0.00295	−0.8	−0.00209	0.00298	−0.7
Seattle	0.00102	0.00096	1.1	0.00479	0.00124	3.9	0.00497	0.00124	4.0
Spokane	0.00230	0.00089	2.6	0.00248	0.00108	2.3	0.00254	0.00110	2.3
Youngstown	0.00104	0.00212	0.5	−0.00268	0.00345	−0.8	−0.00245	0.00346	−0.7
Maximum likelihood	0.00155	0.00036	4.4	0.00236	0.00063	3.7	0.00251	0.00067	3.7

[a] Penalized spline model.

Table A.13. City-Specific and Combined Results of Analysis of PM_{10} Associations with Pneumonia Admissions[a]

City	PM_{10} Lag 0/1			Quadratic Distributed Lag			Unconstrained Distributed Lag		
	β	SE	t	β	SE	t	β	SE	t
Birmingham	0.00000	0.00065	0.0	0.00078	0.00087	0.9	0.00090	0.00087	1.0
Boulder	0.00308	0.00280	1.1	0.00496	0.00425	1.2	0.00505	0.00426	1.2
Canton	0.00046	0.00163	0.3	−0.00305	0.00270	−1.1	−0.00305	0.00271	−1.1
Chicago	0.00295	0.00032	9.3	0.00123	0.00049	2.5	0.00128	0.00049	2.6
Colorado Springs	0.00391	0.00163	2.4	0.00685	0.00257	2.7	0.00660	0.00257	2.6
Detroit	0.00223	0.00036	6.1	0.00075	0.00059	1.3	0.00171	0.00059	2.9
Minneapolis	0.00191	0.00078	2.4	0.00069	0.00116	0.6	0.00061	0.00117	0.5
Nashville	−0.00098	0.00124	−0.8	0.00024	0.00198	0.1	0.00022	0.00199	0.1
New Haven	0.00217	0.00100	2.2	−0.00225	0.00156	−1.4	−0.00221	0.00156	−1.4
Pittsburgh	0.00044	0.00041	1.1	−0.00037	0.00062	−0.6	−0.00042	0.00062	−0.7
Provo/Orem	0.00009	0.00117	0.1	0.00034	0.00128	0.3	0.00068	0.00129	0.5
Seattle	0.00117	0.00060	1.9	0.00113	0.00079	1.4	0.00099	0.00080	1.2
Spokane	0.00010	0.00060	0.2	0.00098	0.00071	1.4	0.00114	0.00072	1.6
Youngstown	0.00030	0.00157	0.2	−0.00100	0.00253	−0.4	−0.00081	0.00255	−0.3
Maximum likelihood	0.00123	0.00038	3.3	0.00064	0.00031	2.1	0.00080	0.00033	2.4

[a] Penalized spline model.

Table A.14. Results of Distributed Lag Model for Daily Deaths and PM_{10} in 10 US Cities[a]

City	GAM with Stricter Criteria		Natural Spline Model		Penalized Spline Model	
	Quadratic	Unconstrained	Quadratic	Unconstrained	Quadratic	Unconstrained
Birmingham	0.000311 (0.000401)	0.000299 (0.000539)	0.000042 (0.000432)	0.000299 (0.001134)	−0.000654 (0.000518)	−0.000645 (0.000717)
Canton	0.001595 (0.001015)	0.001675 (0.001353)	0.001207 (0.001164)	0.001267 (0.001567)	0.001681 (0.001116)	0.001792 (0.001513)
Chicago	0.000936 (0.000190)	0.000874 (0.000265)	0.000965 (0.000207)	0.000899 (0.000285)	0.001011 (0.000185)	0.000978 (0.000261)
Colorado Springs	0.001864 (0.001043)	0.001674 (0.001252)	0.001809 (0.001134)	0.001571 (0.001431)	0.001786 (0.001105)	0.001678 (0.001397)
Detroit	0.001549 (0.000216)	0.001523 (0.000302)	0.001311 (0.000261)	0.001309 (0.000362)	0.001150 (0.000251)	0.001139 (0.000354)
Minneapolis	0.001930 (0.000381)	0.001867 (0.000511)	0.001669 (0.000416)	0.001614 (0.000549)	0.002101 (0.000404)	0.002049 (0.000545)
New Haven	0.001819 (0.000570)	0.001778 (0.000771)	0.001889 (0.000661)	0.001858 (0.000873)	0.001876 (0.000638)	0.001778 (0.000848)
Pittsburgh	0.000752 (0.000224)	0.000712 (0.000300)	0.000200 (0.000258)	0.000169 (0.000352)	0.000577 (0.000243)	0.000561 (0.000334)
Seattle	0.001494 (0.000283)	0.001490 (0.000337)	0.001249 (0.000314)	0.001245 (0.000387)	0.001370 (0.000298)	0.001372 (0.000364)
Spokane	0.000687 (0.000456)	0.000631 (0.000487)	0.000933 (0.000475)	0.000863 (0.000514)	0.000958 (0.000569)	0.000944 (0.000609)
Overall (random effects)	0.001189 (0.000166)	0.001124 (0.000161)	0.000997 (0.000187)	0.000975 (0.000158)	0.001039 (0.000105)	0.001033 (0.000142)

[a] All data are β ± SE.

APPENDIX B. Particulate Air Pollution and Daily Deaths: A Multicity Case-Crossover Analysis (Author: Joel Schwartz)*

ABSTRACT

Numerous studies have reported that day-to-day changes in particulate air pollution are associated with daily deaths. Recently, several reports have indicated that the software used to control for season and weather in some of those studies had deficiencies, and suggested the use of regression splines as an alternative. In this study, I propose the use of the case-crossover design as an alternative. This approach controls for seasonal patterns by design instead of by complex modeling, and, by construction, controls for all slowly varying covariates (age, smoking, etc) by matching. I applied that approach to a study of 10 US cities. Weather and day of the week were controlled for in the regression.

I found that a 10 µg/m^3 increase in PM_{10} was associated with a 0.35% increase in daily deaths (95% CI, 0.21, 0.49). The association appeared quite linear. These results are similar to results of the NMMAPS study when using 3.5 *df*/year to control for season but are larger than the results using 7 *df*/year to control for season. Thus the association between PM_{10} and daily deaths persisted across different modeling strategies.

INTRODUCTION

The case-crossover design, introduced to epidemiologic research by Maclure (1991), is an attractive method for investigating the acute effects of an exposure. For example, the method has been used to investigate triggers of myocardial infarction (Mittleman et al 1993). In recent years, it has been applied to analysis of acute effects of environmental exposures, especially air pollution (Lee and Schwartz 1992; Neas et al 1999; Sunyer et al 2000; Levy et al 2001). In the case-crossover approach, a case-control study is conducted whereby each person who had an event is matched with herself on a nearby time period where she did not have the event. The subject's characteristics and exposures at the time of the case event are compared with those of a control period in which the event did not occur. Each risk set consists of one individual as that individual crosses over between different exposure levels in the interval between the two time periods. These matched pairs may be analyzed using conditional logistic regression. Multiple control periods may be used.

The data are then analyzed as a matched case-control study. Applied to the association of air pollution with risk of death, the approach has several advantages. First, it clarifies a key feature of the study of acute response to air pollution. Because in this analysis each subject serves as her own control, use of a nearby day as the control period means that all covariates that change slowly over time (such as smoking history, age, body mass index, usual diet, diabetes mellitus) are controlled for by matching.

The second advantage involves the method of control for seasonal variations in mortality risk. The other possible technique for analyzing the association of day-to-day changes in air quality with day-to-day changes in daily deaths or hospital admissions is the Poisson regression of daily data (Schwartz and Dockery 1992a). Because these regressions make comparisons across the full range of data, including multiple years, the calculations must control for season and long-term time trends. While several approaches have been taken to control for seasonal patterns in the data (Schwartz and Dockery 1992b; Katsouyanni et al 1997), GAMs were applied to these analyses in 1994 (Hastie and Tibshirani 1990; Schwartz 1993). That technique quickly became the standard in subsequent studies (Samet et al 2000; Katsouyanni et al 2001). These models are attractive because they use smooth curves to control for season (and weather).

Nonparametric smoothing is attractive when one believes a nonlinear association exists with a covariate because the method is more flexible than parametric approaches (Hastie and Tibshirani 1990). Recent studies, however, have reported problems with the algorithms that implement these models (Dominici at al 2002). In individual studies, the standard errors of the parametric terms, including the hypothesis variables, are not correctly estimated (Ramsay et al 2003). Because underestimates of within-location standard errors lead to larger estimates of the between-location standard errors, recent reanalyses of multilocation studies show no evidence of bias in estimated standard errors of combined-effect estimates (Dominici 2002; Schwartz 2002).

These problems have encouraged reanalyzing previous studies to confirm whether the reported associations still hold. Natural splines are a possible alternative and have previously been used (Schwartz 1993), but they have some sensitivity to the knot locations for the splines. Natural splines have recently been applied to a reanalysis of the National Mortality and Morbidity Air Pollution Study (NMMAPS), a multicity study of particulate air pollution and daily deaths (Dominici et al 2002). There has been a continuing debate over how many degrees of freedom are appropriate to control for season, without overcontrol, and

switching from nonparametric smoothing to natural splines does not resolve that issue.

Results from the large number of studies concerning daily changes in air pollution and deaths have formed the basis for tighter air pollution standards in both the United States and Europe. Recent questions over the reliability of these estimates have, therefore, considerable importance for public health policies.

The case-crossover design controls for seasonal variation, time trends, and slowly varying time covariates by separating the case and control periods in each risk set by a relatively small time interval. Bateson and Schwartz (1999, 2001) demonstrated that by choosing control days close to event days, even strong confounding of exposure by seasonal patterns could be controlled in this approach. Also, because the strata of days are matched for each individual, it is straightforward to combine events from multiple locations in a single analysis. The difference in seasonal patterns from city to city has prevented this approach in multicity studies using Poisson regression. This makes the approach an attractive alternative to the Poisson models. While Bateson and Schwartz (2001) have shown that the power is lower in the case-crossover approach, this is less of a concern in a large multicity study.

While sampling control days in a manner that removes seasonal confounding is straightforward, subtle selection bias can occur in these analyses (Levy et al 2001). Days before the first event serve as control days but cannot serve as event days, and occasional days with missing data for exposure during the event series can further increase selection bias. Several approaches have been used to address this problem, and this paper uses the approach of Bateson and Schwartz (2001), which calculates and subtracts the bias. I have applied this approach to a multicity study of particulate air pollution and daily deaths in a study of ten US cities that have previously been analyzed using smoothing techniques.

DATA AND METHODS

Most cities in the United States only monitored PM_{10} once every six days. I focused on 10 US cities with daily data to obtain adequate power. They were Canton, Ohio; Birmingham, Alabama; Chicago, Illinois; Colorado Springs, Colorado; Detroit, Michigan; Minneapolis, Minnesota; New Haven, Connecticut; Pittsburgh, Pennsylvania; Seattle, Washington; and Spokane, Washington. I chose the metropolitan county containing each city except for Minneapolis, which was combined with St Paul and analyzed as one city.

Daily Mortality

Daily deaths in the county containing each city were extracted from tapes prepared by the National Center for Health Statistics (NCHS) for calendar years 1986 through 1993. Deaths from accidental causes (ICD 9 $\geq$ 800) were excluded, as were all deaths that occurred outside of the city. Daily measurements of mean temperature and relative humidity were obtained from the nearest National Weather Service Surface Station (EarthInfo CD-rom of National Climatic Data Center Surface Airways, EarthInfo, Boulder CO).

Air pollution data for PM_{10} were obtained from the EPA's Aerometric Information Retrieval System (AIRS). Many of the cities have more than one monitoring location, requiring a method to average data from multiple locations. This study used an algorithm as previously reported (Schwartz 2000). To ensure that our exposure measure represented the general population exposure and not local conditions affecting only the immediate vicinity of a given monitor, the correlations among all monitors in each county were computed. Monitors within the lowest tenth percentile of the correlation across all counties were excluded. Some monitors only measure PM_{10} one day in six, and different monitors have different means and standard deviations. I did not want the daily pollution value to change from one day to another because different monitors were reported, but I wanted to capture any differences in actual ambient levels.

In each city the daily mean among monitors for each pollutant was calculated using an algorithm that accounted for differences in the annual mean and standardized deviations of each monitor as follows: 1) The daily standardized deviations were averaged for each monitor on each day. 2) These averages were then multiplied by the standard deviation of all of the monitor readings for the entire year and added back in the annual average of all of the monitors (Samet et al 2000). I used the air pollution concentration the day before each death as the exposure variable because the NMMAPS (Dominici et al 2002) study found that to be the most predictive single-day exposure. The NMMAPS 90-cities study could not analyze multiday exposures because most of their cities only measured PM_{10} one day in six. Use of the same exposure metric facilitated comparisons between the studies.

Analytic Strategy

Two analyses were conducted. Both used conditional logistic regression to analyze the data in a case-crossover design. Matched strata were constructed for each subject: the event day (day of death) and 18 matched control days. These days were chosen to be days 7 to 15 before the event day, and days 7 to 14 after the event day. We chose control days far

Table B.1. City-Specific or Descriptive Statistics for Each of the 10 Cities

City	Time Period	Temperature (°C)[a]	Relative Humidity[a]	PM_{10} (µg/m^3)[a]
Birmingham	4/1/87–12/31/93	11, 18, 24	62, 71, 80	20, 31, 46
Canton	1/1/89–12/24/94	2, 11, 19	66, 74, 82	19, 26, 34
Chicago	3/1/88–12/24/94	2, 11, 19	62, 70, 79	23, 33, 46
Colorado Springs	7/1/87–12/24/94	2, 11, 18	39, 51, 66	18, 23, 31
Detroit	5/1/86–12/24/94	2, 11, 19	64, 71, 79	21, 32, 49
Minneapolis	4/1/87–12/24/94	−1, 9, 19	60, 69, 78	17, 24, 35
New Haven	5/1/87–12/31/91	3, 12, 20	57, 67, 77	17, 26, 38
Pittsburgh	1/1/87–12/24/94	3, 12, 20	61, 70, 79	19, 30, 47
Seattle	1/1/86–12/24/94	7, 11, 16	67, 77, 85	18, 27, 39
Spokane	10/1/85–12/24/94	2, 8, 16	49, 68, 84	23, 36, 57

[a] Each set of 3 numbers refers to the 25th, 50th, and 75th percentiles.

from the event day to avoid serial correlation in the pollution and mortality data. Control days were symmetrically chosen about the event day because Bateson and Schwartz (1999) demonstrated that symmetric control days were needed to control for long-term time trends (if present). Navidi (1998) pointed out that bidirectional sampling is needed to avoid some biases in the case-control method and does not present any conceptual difficulties as long as death of the subject does not affect the air pollution concentrations.

In all analyses, I controlled for day of the week, temperature and relative humidity. Temperature may be nonlinearly related to deaths, and so I used regression splines to control for temperature on the day of death and the day before death. These splines used 3 *df* each. Relative humidity was similarly controlled for.

The first analysis used a two-stage approach. A city-specific regression was fit using the matched strata from each city. The log odds ratios from those 10 analyses were then combined using the iterative maximum likelihood algorithm of Berkey and coworkers (1995). In this analysis, the splines for temperature could have different coefficients in different studies. Because the control days are chosen close to the event day in the case-crossover analysis, the range of temperature variation, and the range of its effects, is lower than in other study designs. This suggested that we could aggregate the strata and analyze the association with air pollution across all ten cities in a single model. This is an attractive capability of this study design and constituted our second analysis. We compared the results from the two analyses to determine how sensitive the results were to modeling weather with one set of splines for all locations.

RESULTS

Table B.1 shows the 25th, 50th, and 75th percentiles of the distribution of daily deaths, PM_{10} levels, temperature, and relative humidity in each of the ten locations. Weather was only modestly correlated with PM_{10} in these locations (Table B.2).

In the two-stage analysis, I found a significant association between PM_{10} and daily deaths. The magnitude of the association was a 0.35% increase in daily deaths per 10 µg/m^3 increment of PM_{10} (95% CI, 0.21, 0.49). Combining the strata and analyzing in one stage had little impact on the estimate (0.34% increase, 95% CI, 0.21, 0.48). The individual city results are shown in Table B.3. There was no evidence for heterogeneity in the association (χ^2 = 4.16 on 9 *df*, P = 0.90).

To test the shape of the dose-response relation, I replaced the linear term for PM_{10} with the indicator variables for days when concentrations were between 15 and 25 µg/m^3, between 25 and 34 µg/m^3, between 35 and

Table B.2. Correlation Between PM_{10} and Other Environmental Variables in 10 Cities

City	Temperature	Relative Humidity
Birmingham	0.26	−0.30
Canton	0.42	−0.16
Chicago	0.36	−0.30
Colorado Springs	−0.34	−0.11
Detroit	0.37	−0.14
Minneapolis	0.29	−0.35
New Haven	0.05	−0.15
Pittsburgh	0.45	−0.23
Seattle	−0.22	−0.11
Spokane	−0.01	−0.19

Table B.3. Percentage of Change in Daily Deaths for 10 µg/m³ Increase in PM_{10} by City, and Overall, Using Either a Two-Stage or Single-Stage Analysis

City	% Change	95% CI
Birmingham	−0.01	−0.58, 0.56
Canton	−0.02	−1.67, 1.65
Chicago	0.39	0.14, 0.64
Colorado Springs	0.35	−1.01, 1.74
Detroit	0.47	0.15, 0.78
Minneapolis	0.63	0.06, 1.21
New Haven	0.27	−0.63, 1.19
Pittsburgh	0.30	−0.05, 0.65
Seattle	0.14	−0.32, 0.61
Spokane	0.32	−0.20, 1.21
All two-stage	0.35	0.21, 0.49
All single-stage	0.34	0.21, 0.48

44 µg/m³, and 45 µg/m³ and above. The days with concentrations below 15 µg/m³ served as the reference level. Figure B.1 shows the results.

DISCUSSION

This analysis confirms the recent report from NMMAPS, that when using parametric regression techniques to avoid the problems in current GAM software, a significant association is still evident between daily PM_{10} levels and daily deaths. The new analyses extend that finding by a) using a very different approach, indicating a robustness of the results to type of modeling, and b) using an approach that avoids modeling seasonal patterns, obviating arguments about the complexity of those models.

The magnitude of the association for the 10 cities was larger than the central estimate reported for 90 cities in the NMMAPS study using natural splines (0.35% vs 0.21%). This difference could be due to sampling variability, to the choice of cities studied, or to the difference in the methods. If it is due to the latter, the difference is most likely the question of degree of freedom rather than the use of natural splines per se because any sensitivity to knot location should average out in a multicity study. The NMMAPS study used a range of degrees of freedom to control for season but centered on 7 *df*/year. Most other studies of air pollution and daily deaths have used 3 to 4 *df*/year (Goldberg et al 2001; Katsouyanni et al 2001). In one of their sensitivity analyses, Dominici and coworkers (2002) reported the effect of using 3.5 *df*/year in their models. The effect estimate for PM_{10} increased and became indistinguishable from the results in this study.
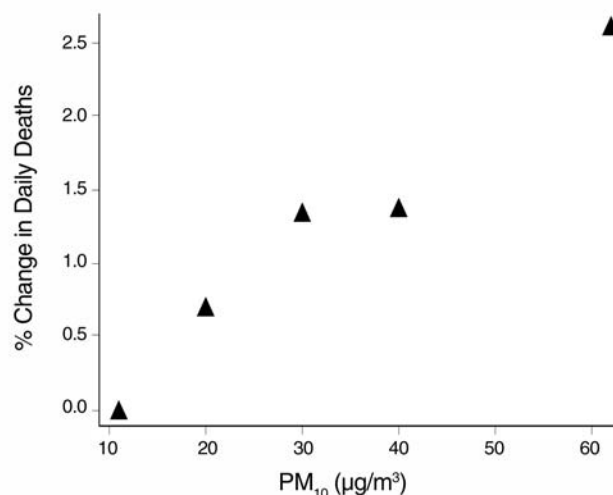
Figure B.1. Dose response between PM_{10} and daily deaths. Percentage increase in daily deaths on days with PM_{10} concentrations in the ranges of 15–24 µg/m³, 25–34 µg/m³, 35–44 µg/m³, and 45 µg/m³ and greater, compared to reference days when concentrations were below 15 µg/m³.

Because simulation studies have demonstrated that the control sampling strategy used in this analysis is adequate to control for strong seasonal confounding, the lower degree of freedom appears to be more appropriate for spline models. The problems associated with overcontrol of season have been described (Schwartz et al, elsewhere in this report).

Like the NMMAPS analysis, this study focused on the association of death with PM_{10} concentrations on the previous day. The choice of a one-day lag is not arbitrary, however. Braga and coworkers (2001), in a Poisson regression analysis of these same 10 cities, fit models that simultaneously included PM_{10} concentrations on the day of death and on the previous five days. These models were fit for specific causes of death. The use of 6 correlated exposure measures in the model led to imprecise estimates within each city. Those estimates are unbiased, however, and by combining estimates across the 10 cities in a combined analysis, interesting patterns emerged.

For deaths from myocardial infarction, most of the effect was for pollution on the day of death, with some effect at lag 1, and nothing at longer lags. This is consistent with the growing evidence that myocardial infarctions can be acutely triggered by stressors (Mittleman et al 1993). Indeed, a recent case-crossover analysis reported that increases in particulate air pollution concentrations were associated with an increased risk of death within hours of exposure (Peters et al 2001).

For all cardiovascular deaths, the association was more evenly split between lags 0 and 1 of pollution, again falling to zero after that. This suggests that other cardiovascular events occur with a bit more lag. For deaths due to pneumonia and COPD, there was no effect at lag 0. The effect was limited to exposures 1 and 2 days before the event. This is consistent with observed associations between particle exposures and increases in inflammation and with observations of Zelikoff and coworkers (1999). They infected rats with *Streptococcus pneumoniae* and then exposed them to concentrated air particles. The infected rats had twice the bacterial burden and area of the lung involved compared to rats breathing filtered air. The full effect, was not seen for 48 hours, however.

Hence, a lag of one day between exposure and death represents the day of overlap between the different lags associated with different causes of death. This overlap makes it the most appropriate choice for a single-day model of all-cause mortality. Because of the one-day-in-six sampling frame for EPA monitors, the NMMAPS 90-cities study was limited to choosing a single day.

The study of Braga and coworkers (2001) illustrates a limitation of the case-crossover approach. Because of serial correlation in exposure, if we wanted to simultaneously examine the effects of multiple lags, we would have to push our control dates further from the event date. As Bateson and Schwartz (2001) have shown, that would entail some risk of bias. Moreover, the reduced power (80% relative to a Poisson model) of our approach would also start to become more important in a model with six correlated exposure variables. While this is a limitation of the case-crossover approach, its advantages are that it controls for season without using a complex model is more familiar to epidemiologists, and easily allows the matched strata to be combined across multiple studies. Using this approach, I have confirmed that short-term changes in PM_{10} are associated with short-term changes in daily deaths.

ABOUT THE AUTHORS

Joel Schwartz, associate professor of environmental epidemiology at Harvard School of Public Health and Harvard Medical School, received his doctorate from Brandeis University in theoretical physics and is a MacArthur Fellow. Schwartz has held appointments as a visiting scientist at the University of Basel, Switzerland, and at the University of Wuppertal, Germany, and has served as a Senior Scientist at the US EPA. His research focuses on the health effects of air and water pollution, lead, and meteorologic questions regarding the modeling of covariates in epidemiologic studies.

Antonella Zanobetti, research associate in environmental epidemiology at Harvard School of Public Health, received her doctoral degree from the University of Florence. Her fields of research include statistical analysis in environmental epidemiology, with particular interest in the short-term effect of air pollution on mortality and morbidity, and air pollution effects on cardiovascular health.

ABBREVIATIONS AND OTHER TERMS

APHEA	Air Pollution and Health: A European Approach
β	coefficient
CI	confidence interval
COPD	chronic obstructive pulmonary disease
CVD	cardiovascular disease
df	degree of freedom
EPA	Environmental Protection Agency (US)
GAM	generalized additive model
gam	GAM function in S-Plus software
GLM	generalized linear model
ICD 9	*International Classification of Diseases*, 9th Revision
LOESS	locally weighted smoothing function
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
ns	natural spline
PM	particulate matter
PM_{10}	PM < 10 μm in aerodynamic diameter
REML	restricted maximum likelihood
SE	standard error
t	ratio of the coefficient to the standard error

Commentary on Revised Analyses of NMMAPS II Data

A Special Panel of the Health Review Committee

BACKGROUND

Study of the short-term effects of air pollution has relied heavily on observational findings from time-series studies. In these studies, the association between short-term changes in air pollution and adverse health outcomes is typically assessed using measurements of air pollutant concentrations and corresponding counts of deaths or hospital admissions. To minimize spurious, or biased, associations between these time-varying measures, various analytic methods are used to account for effects of concurrent factors varying over time that correlate with pollutant concentrations and with daily counts of mortality and morbidity. When the impact of these other factors are not controlled, the effects attributed to air pollution may be actually due to these other factors, including weather (typically temperature and relative humidity), as well as unmeasured factors that also vary with time.

If unmeasured factors change slowly, they are reflected in longer time trends than those that appear to be relevant to short-term air pollution effects. While many analytic methods are available to control for such factors in time-series analyses, the use of generalized additive models (GAMs*) has been preferred since 1996. GAMs allow for smooth functions of time to capture the slower fluctuations in mortality or morbidity. They are also used to allow for effects of some measured variables, such as temperature, without the need to assume a specific parametric model for the effect.

---

* A list of abbreviations and other terms appears at the end of the Commentary.

Footnotes throughout this Commentary indicate sources in either of the two revised analyses of NMMAPS II data (Dominici et al or Schwartz et al) or in the original NMMAPS II Research Report 94 (Samet et al 2000).

During the review process, the Special Panel of the Health Review Committee and the investigators had the opportunity to exchange comments and to clarify issues in both the reports and in the Commentary.

This document has not been reviewed by public or private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views of these parties, and no endorsements by them should be inferred.

Problems with the GAM module in commonly used S-Plus statistical analysis software when applied to time-series studies of air pollution came to light in May 2002. First, Johns Hopkins investigators who were conducting sensitivity analyses on the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) dataset found that default convergence criteria used in fitting a GAM could lead to miscalculation of the estimates of effect (Dominici et al 2002). Specifically, the convergence criteria preprogrammed in the GAM module of S-Plus prematurely terminated the iterative process used in estimating the regression coefficients. Further, investigators at Health Canada independently became aware that a programming shortcut used in calculating the standard error for regression coefficients underestimated the true standard error (Ramsay et al 2003). Dominici and colleagues (2002) suggested that the net effect of these problems was typically to overestimate the effect estimate and to underestimate its standard error.

NMMAPS is arguably the most influential of time-series studies of short-term changes in air pollutant concentrations and the associations with mortality and hospitalizations. This study was carried out by two collaborating groups of investigators. Air pollution and mortality in the 90 largest cities in the United States was studied by investigators from Johns Hopkins University. They made use of available air pollutant monitoring data to carry out city-specific analyses in a standard manner. The resulting estimates of city-specific effects were then combined using hierarchical Bayesian methods to estimate effects for regions of the country and for the country as a whole. Investigators at Harvard University studied pollution and both hospitalizations and mortality in smaller groups of 14 and 10 cities, respectively (selected because of data available on daily air pollution concentration and hospital admissions). Both parts of NMMAPS made extensive use of GAMs to control for time trends and weather factors (Samet et al 2000).

When the NMMAPS investigators identified problems with applying GAMs, they responded rapidly by conducting a series of revised analyses and sensitivity analyses to address both the convergence criteria and the standard error issues. Their approaches to uncovering the errors and attempting to correct them are detailed in the accompanying reports. This Commentary summarizes the methods used to identify and quantify the problems and the results in both parts of the study. An assessment of the methods and revised findings is followed by conclusions and recommendations of the Special Panel and thoughts on the impact of problems associated with applying GAM software.

SUMMARY OF REPORTS

ANALYTIC METHODS

The investigators' revised analyses addressed both problems with the use of GAMs (convergence and standard error estimation), while purposefully leaving most other aspects of the analyses unchanged. Specifically, Dominici, Schwartz, and colleagues replaced the GAM functions used in the previous analyses as follows:

1. *Identical GAM functions with stricter convergence criteria*. These analyses were designed to correct the GAM convergence problem while acknowledging that the problem with standard error estimates remained in the basic GAM program.
2. *Generalized linear models (GLMs) with natural cubic splines,* using the same degrees of freedom (*df*) as used in the previous GAM functions, rounded to integer values. GLMs with natural cubic splines are fully parametric because they can be expressed in terms of a finite number of parameters representing coefficients of the splines. These models should provide unbiased estimates of the standard errors and should converge correctly. However, these GLMs with natural cubic splines differ in other respects from GAMs: their capacity to match fluctuations in the relations between variables is not identical to that of GAMs for the same number of degrees of freedom, and their form depends somewhat on placement of knots. In the analyses of data from the 90 cities, knots were placed at equally spaced quantiles of the distribution of explanatory variables.

Schwartz and colleagues also used two other methods:

3. *Penalized splines.* Using the same degrees of freedom as in the original GAM functions, penalized splines were used to smooth terms in one step rather than one

covariate at a time until convergence is reached (as in GAM). These curved functions are similar in concept to GAM smooth functions, but they do not require the same approximation to estimate standard errors, which should thus be well estimated.

4. *Case-crossover matching.* At first sight, this analytic design takes a different approach to controlling for confounding. The case-crossover approach seeks to control for fluctuations in mortality over time by matching rather than by modeling the time effect explicitly in a regression model. Specifically, Schwartz and colleagues used this technique to assess mortality from nonexternal causes in 10 US cities for which daily monitoring of pollutants had been conducted. They matched the particulate matter (PM) level on the day before each death to the level on each of a number of control days 7 to 15 days before and after the death day. Using the case-crossover design, Schwartz and colleagues compared two approaches to obtain an estimate of mean effect across all cities: a two-stage calculation of the mean from the effect estimates for all cities; and a one-stage calculation using data from all 10 cities included in a single model.[1]

As in the original report, Dominici and colleagues applied the same model to each of the 90 cities included in the evaluation of daily mortality. That is, the same variables and smoothing functions were used in each city to control for potential confounding, while parameter estimates and fitted smooth functions were allowed to vary from city to city. Schwartz and colleagues conducted the original and the revised analyses fitting a city-specific model to each of the 14 cities included in the analysis of hospital admissions data.

In these revised analyses Dominici and colleagues combined evidence across cities using both the simpler DerSimonian and Laird (1986) and more complex hierarchical Bayesian methods. Their revised analysis report provides results only from the Bayesian methods; a comparison to the simpler method was published earlier (Dominici et al 2002). They also give more information on heterogeneity and on impact of modifying their model for heterogeneity.[2]

Schwartz and colleagues changed the method used for this metaanalysis component, replacing the simple noniterative method of DerSimonian and Laird (1986) by an iterative maximum likelihood approach, as was recommended in the original NMMAPS II Commentary. This method would not change the city-specific estimates but could change the estimates of means over the 14 cities and the

---

1. Schwartz et al, Appendix B.
2. Dominici et al, Table 2, Figure 11, and Appendix A.

test for heterogeneity. In the current report, Schwartz and colleagues present estimates of mean effects and their confidence intervals (CIs) exclusively using random effects models, whereas previously estimates from fixed and random effects models were presented with the fixed effects results highlighted.

RESULTS

Overall revised effect estimates using GAMs with stricter convergence criteria and using GLMs with natural cubic splines were generally lower than those obtained from the original analyses with GAM and default criteria. City-specific estimates obtained with GLM also had wider confidence intervals. Nevertheless, small short-term associations remained between the concentration of PM less than 10 μm in aerodynamic diameter (PM_{10}) and nonexternal mortality and between PM_{10} and mortality and morbidity from cardiovascular and respiratory disorders. Overall, similar but smaller effect estimates were obtained for individual cities[3] and geographic regions.[4] The investigators observed comparable patterns for lags 0, 1 and 2, with the most consistent association found at lag 1 across 90 cities and less pronounced associations for other lags.[5] Among the 10 cities for which daily PM_{10} concentrations were available, the most consistent association with mortality was reported for two-day mean and distributed lags.[6] In both reports results did not change substantially when copollutants were included in the models.[7]

Air Pollution and Daily Mortality in 90 Cities

Results for the 90 cities were based on available PM monitoring data (collected every 6th day in most cities, but nearly daily in 5 cities), 1987 to 1994 and reported as mean change per 10 μg/m^3 increase in PM_{10} concentration. In the current report, Dominici and colleagues present a mean change at lag 1 across all 90 cities of 0.27% (posterior SE = 0.05) for total mortality obtained with GAM with stricter convergence criteria and a mean of 0.21% (posterior SE = 0.06) when obtained with GLM with natural cubic splines. In contrast, the corresponding original estimate at lag 1 was a 0.41% increase in total mortality (posterior SE = 0.05). The mean change in total mortality for lags 0 and 2 also decreased substantially from 0.22% originally to 0.07% (posterior SE = 0.06) for lag 0 and from 0.28% originally to 0.10% (posterior SE = 0.06) for lag 2.[8] The mean estimate for mortality from cardiovascular and respiratory

causes across 90 cities was reported from the revised analysis only, 0.31% (posterior SE = 0.09) for lag 1,[9] 0.20% (posterior SE = 0.09) for lag 2, and as expected, the lowest estimate (0.13% [posterior SE = 0.09]) for lag 0.

Regional patterns within the 88 cities within the contiguous states remained. The effect estimate for the Northeast was the highest and the Upper Midwest estimate was the lowest for total mortality and for mortality from cardiovascular and respiratory diseases.[10] The posterior mean change for all regions was 0.25% (95% posterior interval = 0.03, 0.48) for total mortality and 0.34% (0.10, 0.57) for cardiovascular and respiratory related mortality.[11] Regional posterior means for total mortality were 0.38% for the Northeast (95% posterior interval = 0.01, 0.76) and 0.19% (−18, 0.55) for the Upper Midwest region.[12] Regional posterior means for mortality from cardiovascular and respiratory diseases were 0.50% (0.05, 0.94) for the Northeast and 0.26% (−0.11, 0.63) for the Upper Midwest. Revised posterior means for total mortality for lags 0 to 2 decreased in all regions with the largest decrease in the Northeast.[13]

Dominici and colleagues also conducted sensitivity analyses varying the adjustment for confounding factors. Increasing the degrees of freedom in the smooth functions of time, temperature and dew point within the model decreased the posterior mean for total mortality for lag 1: 0.32% when the degrees of freedom for variables were halved and 0.17% when the degrees of freedom were doubled.[14] A similar pattern was reported originally.[15]

Dominici and colleagues estimated the variability of city-specific effect estimates across cities.[16] They conducted a Bayesian analysis with a two-stage hierarchical model and estimated the marginal posterior distribution of the heterogeneity parameter. They estimated less heterogeneity in the effect of air pollution on mortality among cities than reported earlier. The posterior mean of the between-city standard deviation decreased from 0.112% (95% posterior interval = 0.022, 0.298), obtained with default GAM, to 0.075% with GLM and natural cubic splines (95% posterior interval = 0.022, 0.298).[17] For the 88 cities, city-specific posterior means calculated with two prior models yielded similar mean relative risks from ambient PM_{10}. The models had different assumptions regarding the presence of

3. Schwartz et al, Figures 14 and 15; Tables A.1, A.2, A.7–A.14, B.3.
4. Dominici et al, Figures 6, 7, and 9; Tables A.1 and A.2.
5. Dominici et al, Figures 2, 3, 8, 12, 14–16.
6. Schwartz et al, Table A.14.
7. Dominici et al, Figures 5 and 12–16; Table A.1. Schwartz et al, Figures 12–13.
8. Dominici et al, Figures 2 and 3.
9. Dominici et al, Figure 4.
10. Dominici et al, Figures 6–9.
11. Dominici et al, Table A.1.
12. Dominici et al, Table A.1.
13. Dominici et al, Figure 8; Samet et al 2000, Figure 23, p 26.
14. Dominici et al, Figure 10.
15. Samet et al 2000, Figure A.1, p 53.
16. Dominici et al, Figures 11, A.1 and A.2, Tables A.1 and A.2.
17. Dominici et al, Table 2.

heterogeneity across cities (model A: 0.25% [95% CI = 0.03, 0.47]; model B: 0.22% [0.02, 0.43]).

Dominici and colleagues present revised results for estimates of effect for lags 0, 1 and 2 for ozone, sulfur dioxide and nitrogen dioxide both without adjustment for other pollutants and with adjustment for each of several copollutants, including PM_{10}. These results changed little from those reported originally except for being less precise.[18] Results for ozone by season demonstrated little change in effect estimates in winter and summer, but a slight increase from original estimates across all seasons.[19]

Morbidity Among Elderly in 14 Cities with Daily PM Data

Daily PM monitoring for an extended period between 1985 and 1994 (years for which hospital admission data were also available) were available for 14 of the cities included in the analyses of morbidity. Results for the association of PM_{10} with hospital admissions for specific diagnoses among those 65 years of age and older varied according to diagnosis in the revised analyses. For hospital admissions for cardiovascular disease (CVD), mean effect estimates per 10 µg/m^3 increase in PM_{10} concentration differed little comparing random effects model of GAM with default and stricter convergence criteria, GLM with natural splines, and models with penalized splines: default GAM, b = 0.00106 [SE 0.00020]; strict GAM, b = 0.00111 [SE 0.00014]; GLM, b = 0.00111 [SE 0.00016]; penalized spline, b = 0.00111 [SE 0.00015] when estimating the unconstrained distributed lag.[20] A moderate decrease was found for the unconstrained distributed lag between corresponding estimates for chronic obstructive pulmonary disease (COPD) (default GAM, b = 0.00284 [SE 0.00134]; strict GAM, b = 0.00250 [SE 0.00066]; GLM, b = 0.00209 [SE 0.00074]; penalized spline, b = 0.00251 [SE 0.00067]).[21] The most substantial drop, however, was observed for pneumonia when the data were analyzed with GLM with natural splines (default GAM, b = 0.00205 [SE 0.00057]; strict GAM, b = 0.00160 [SE 0.00034]; GLM, b = 0.00001 [SE 0.00033]; penalized spline, b = 0.00080 [SE 0.00033]).[22]

When restricting the analysis to days with levels below 50 µg/m^3 PM_{10}, differences between the estimates from original and revised analyses were smaller for a two-day mean concentration. Comparing results reported from the random effects model with GAM and default convergence

criteria and GLM with natural spline, changes in estimates of effect were moderate for cardiovascular admissions (default GAM, b = 0.00144 [SE 0.00017]; GLM, b = 0.00131 [SE 0.00028]) and greater for admissions due to pneumonia (default GAM, b = 0.00243 [SE 0.00065]; GLM, b = 0.00105 [SE 0.00051]).[23]

The patterns for distributed lag effect estimates of PM_{10} relative to the correlation of PM_{10} with temperature and relative humidity remained the same for hospital admissions for all diagnoses, showing no significant increasing trend with increasing correlations. The estimates, however, showed little change for COPD and pneumonia but a slight decrease for CVD.[24]

Results of tests for heterogeneity of city-specific estimates for specific diagnoses conducted with natural splines were similar to those reported in the original report.[25] Evidence for heterogeneity appeared somewhat less for hospital admissions for pneumonia in the revised analyses.

Using the same approach as reported originally to control for other pollutants, results comparable to those reported earlier were obtained from analyses with natural spline and penalized spline models for city-specific and mean effect estimates for PM_{10} concentration relative to the ratio of sulfur dioxide and ozone to PM_{10} for CVD and COPD; however, controlling for ozone increased the effect estimate somewhat for hospital admissions for pneumonia.[26]

As in the earlier data analysis, Schwartz and colleagues conducted metaregression analyses using ecologic measures of social factors (such as college education, unemployment, and poverty) to assess effect modification.[27] The effect estimate of ambient PM_{10} was greater among the less educated, the unemployed, and the poor as shown by admissions for CVD when data were analyzed applying models with natural splines or penalized splines. Little effect modification had been observed with the original GAM analyses. Further, each 5 percentage point increase among populations in the above mentioned characteristics (especially unemployment, poverty, and to a lesser extent, nonwhite race and lack of college education) was associated with increases in the effect estimates for CVD hospital admissions per 10 µg/m^3 increase in PM_{10} concentration (unemployment, GLM, 0.78% change in effect estimate [95% CI = 0.42%, 1.14%]; penalized spline, 0.45% [95% CI = 0.01%, 0.88%]). Changes in effect estimates for COPD and pneumonia relative to social factors were less consistent.[28]

18. Dominici et al, Figures 5, 12, and 14–16; Samet et al 2000, Figures 24-29, pp 27-28.

19. Dominici et al, Figure 13; Samet et al 2000, Figure 30, p 28.

20. Schwartz et al, Tables A.1, A.7, and A.11; Samet et al 2000, Table D.5, p 67.

21. Schwartz et al, Tables A.2, A.8, and A.12; Samet et al 2000, Table D.7, p 68.

22. Schwartz et al, Tables A.3, A.9, and A.13; Samet et al 2000, Table D.6, p 67.

23. Schwartz et al, Table A.10; Samet et al 2000, Table D.8, p 68.

24. Schwartz et al, Figure 11; Samet et al 2000, Figure 32, p 40.

25. Schwartz et al, Table 4; Samet et al 2000, Table 35, p 38.

26. Schwartz et al, Figures 12–13; Samet et al 2000, Figure 33, p 41.

27. Samet et al 2000, pp 34, 35, 38, 40.

28. Schwartz et al, Table 6; Samet et al 2000, Table 16, p 38.

Mortality Among Elderly in 10 Cities with Daily PM Data

Ten cities with daily PM_{10} monitoring were included in the original analyses of mortality data.[29] Schwartz and colleagues conducted revised analyses with GAM with stricter convergence criteria, natural spline models, and penalized spline models. Neither of the latter methods is affected by biased estimation of standard errors nor the iterative loop (backfitting algorithm) associated with the GAM function. The penalized spline method provides greater flexibility in smoothing and depends less on arbitrary choice of knot points. The mean effect estimates for unconstrained distributed lags decreased from the original 1.3% (95% CI = 1.0, 1.5) increase in daily deaths per 10 µg/m^3 increase in PM_{10} to 1.1% (95% CI = 0.8, 1.4) with stricter GAM and to 1.0% with penalized splines (95% CI = 0.8, 1.3) and GLM (95% CI = 0.7, 1.3).[30]

A case-crossover matched analysis compared revised results for mortality to those obtained with another analytic method.[31] A two-stage analysis combined city-specific effect estimates to obtain a common effect estimate across cities, and a single-stage analysis assumed no heterogeneity of effect among cities and combined matched sets from all cities. The two-stage analysis estimated a 0.35% (95% CI = 0.21, 0.49) increase in daily deaths per 10 µg/m^3 increase in PM_{10}; results obtained from the single-stage analysis [0.34% (95% CI = 0.21, 0.48)] were similar. (Results for unconstrained distributed lags in the time-series analyses were substantially larger.) To explore the dose-response relation between PM_{10} and mortality, Schwartz and colleagues also compared the effect of different ambient levels to an ambient level below 15 µg/m^3. Compared with daily deaths observed at this concentration, the percentage of increase in daily deaths rose by 0.7 at PM_{10} concentrations of 20 µg/m^3 to approximately 1.2 to 1.3 at PM_{10} concentrations of 30 and 40 µg/m^3 and to 2.6 at concentrations higher than 60 µg/m^3.

Schwartz and colleagues also presented results of other sensitivity analyses and simulations to evaluate the effect of varying adjustment for confounding factors.[32]

DISCUSSION OF ANALYTIC METHODS

The revised analyses of the NMMAPS data set provide regression effect estimates and their variability, specifically the standard errors of these estimates, using different models and software programming. In addition,

---

29. Samet et al 2000, pp 54–61.
30. Schwartz et al, Tables 45 and A.14; Samet et al 2000, Table B.3, Figure B.3.
31. Schwartz et al, Appendix B.
32. Schwartz et al, Figures 1–10.

the sensitivity of the findings to modeling approaches and to choice of parameters was explored.

Familiarity with modern statistical methods is needed to fully appreciate the necessity for these revised analyses and the grounds for uncertainty in interpreting the findings. Thus this section attempts to provide a summary understanding of the issues. The Detailed Discussion of Analytic Methods section is for those who wish to understand more fully the details underlying this summary and our conclusions.

Time is a surrogate measure for unknown or unmeasured factors that affect daily mortality or morbidity (eg, hospitalizations) and might also be associated with pollution. When such time-varying factors are not taken into account, any apparent associations with short-term changes in air pollutant concentrations could be partly or entirely due to these factors. Unfortunately, if control is too strict, which in this case consists in modeling time effects too finely, the estimate of air pollution effect can become imprecise and part of a true pollution effect can be absorbed into the time effect. This has been described as *undersmoothing* and corresponds to using too many degrees of freedom in modeling time. Alternatively, if control is not strict enough, effects due to other time-varying causes may be incorrectly attributed to air pollution. This can be described as *oversmoothing* and corresponds to using too few degrees of freedom in modeling time. One element of uncertainty is due to the fact that control for these other time-varying factors can never be complete. The smaller the effects, and all would agree that air pollution effects are small, the better the control for time needs to be. Effects of weather (typically temperature and relative humidity in these studies) have similar potential pitfalls when modeled in a fashion similar to time.

The most common approach is to model time explicitly, as with Poisson regression models that are ubiquitous in air pollution time-series studies. Many methods can be used to model time in these models. Examples in the NMMAPS analyses include GAMs with locally weighted smooth (LOESS) functions, GLMs with natural cubic spline functions, and penalized splines. Time can also be modeled implicitly, as in the case-crossover approach used as a sensitivity analysis for assessment of mortality. At first glance, the case-crossover approach appears radically different from the other approaches used; however, as discussed in the Detailed Discussion of Analytic Methods, it may present similar problems.

In general, approaches to determining the optimal degree of control for time (degree of smoothing or degrees of freedom) have often relied on measures of how well the modeling fits the temporal pattern. The Akaike information criterion (AIC) has been used as one such measure, but measures of goodness of fit in this setting do not help to determine the appropriate degree of control for time. Another

DETAILED DISCUSSION OF ANALYTIC METHODS

ANALYTIC METHODS FOR MODELING CALENDAR TIME AND AIR POLLUTION EFFECTS WITHIN CITIES

Both sets of investigators carried out the statistical analyses in two stages. In the first stage, each city was analyzed independently of the others to obtain a vector of parameter estimates, β, together with its variance-covariance matrix. The second stage, effectively a metaanalysis, used these estimates as data and focused on exploring heterogeneity of estimates and on estimating a single value. This section of the discussion concentrates on a specific aspect of modeling calendar time, which is the component of the analysis most affected by the problems with applying GAMs.

NEED TO INCORPORATE CALENDAR TIME IN ANALYSES

In analyses of daily air pollution and mortality data, calendar time itself is not a causal variable; calendar time is a surrogate for other unmeasured (or unmodeled) variables that may have causal effects. Three distinct types of such causal variables should be distinguished:

- *unmeasured confounders*, variables which have a causal effect on mortality and which vary in time in a similar manner to air pollution,
- *causal variables*, which vary systematically over time but in a manner unrelated to air pollution, and
- *causal aspects of air pollution itself*, imperfectly captured by available measurements.

The first two variables necessitate some attempt to allow for calendar time effects. First, if some statistical approach can be used to control for calendar time, this will to some extent control for unmeasured confounders (although such control cannot be perfect). Second, even when they do not confound air pollution effects, other variables may induce autocorrelations in mortality and morbidity time series. Failure to take account of these can lead to incorrect standard errors of effect estimates and incorrect hypothesis tests. However, the third source of variation of mortality with time—that due to unmeasured aspects of air pollution itself—can lead to a negative consequence of controlling for calendar time. In this case, controlling for time removes some of the causal effect of air pollution in the analysis. Unfortunately, however, if we wish to address the possibility of confounding, this is a price that must be paid as part of a complete analysis of the sensitivity of findings to differing levels of control.

In the current analyses, control for calendar time was carried out in one of two ways. In parametric and semiparametric regression models, the time effect was incorporated in the model as a smooth function. In contrast, the case-crossover analysis attempted to eliminate time effects by matched choice of control days. In the former approach, the effect of time is modeled explicitly while, in the latter, the model is implicit. The aim of both approaches is to simulate, in analysis, a natural experiment in which air pollution level is varied while calendar time is, as far as possible, held constant, in the hope that one will also control for other unmeasured confounders. Clearly, however, calendar time cannot be held constant because to do so would also hold the level of air pollution constant. Instead we must control time within bounds and focus on short-term variations in air pollution levels. This adjustment can never represent complete control for confounding; only confounders which vary more slowly than air pollution can be dealt with, and even these are dealt with imperfectly. The judgment about how closely time should be controlled is ultimately a scientific one and cannot be made on purely statistical grounds.

Analytic methods may differ in three respects:

1. the functional forms used to model time (and temperature) effects (natural cubic splines, local regression smoothers, etc),
2. the approach to statistical inference (maximum likelihood, conditional likelihood, Bayesian methods, etc), and
3. the numerical method for computing estimates and standard errors of estimates.

In the following sections we try to clarify some of the similarities and differences of approach. Most of this discussion focuses on the problem of modeling time effects, but much of it applies to the important issue of modeling temperature effects.

STRATIFICATION BY TIME

The simplest approach to control for confounding by time is stratification. Although this approach has not been used in either of the current analyses, the approach is useful for introducing some issues that affect more complicated approaches to the analysis.

Consider the total study period as being divided into a large number of distinct intervals (strata), each consisting of a number of consecutive days. A simple approach is to fit, in a Poisson regression model, a stratum effect common to all days within the same stratum. Then the effect of calendar time is modeled as a step function. For the purposes of inference about air pollution effects, the time stratum effects are so-called nuisance parameters.

There are two well-known approaches to inference with this model:

- estimation of calendar time effects by maximum likelihood using the Poisson regression approach, or
- elimination of these nuisance parameters by a conditional argument. Here we argue conditionally upon the time stratum within which each death occurred, basing inference on the conditional probability of the day at which death occurred within the stratum.

These approaches are known to lead to identical large-sample results. The latter approach yields the same likelihood as conditional logistic regression of a matched case-control study. It is formally equivalent to a case-crossover analysis in which the case day is the day in which the death occurred and all other days within the stratum are control days. For estimating the remaining parameters in the model, including air pollution effects, the case-crossover analysis is equivalent to Poisson regression analysis and avoids the estimation of a potentially large number of nuisance parameters—the time stratum effects.

CHOICE OF STRATIFICATION INTERVAL AND AUTOCORRELATION

The stratification approach provides some insight into the important question of how finely the time effect should be modeled. If the time strata are wide (a few degrees of freedom per year), the effects of those confounders for which time is a surrogate will be imperfectly captured, resulting in residual confounding. Failure of the step function to capture the time effect may be manifest as positive autocorrelation in the residuals. This autocorrelation violates the assumption of independence of errors made in the Poisson regression model and in the closely related conditional logistic model. This violation in turn may lead to incorrect standard errors in addition to that of bias due to residual confounding.

As the stratification interval is reduced (more degrees of freedom per year), we control for confounding by influences varying on progressively shorter time scales. The finer the stratification, the more closely we would expect to control for confounding. Note that fine stratification will inevitably lead to negative autocorrelation of residuals of the observed death rates from values fitted by the model. However, the penalty to be paid is increased standard error of the estimates of pollution effects; in the language of the case-crossover analysis, the control days are overmatched, leading to little or no variation between case and control days.

A further consequence of very fine stratification is that impact on effect estimates of deficiencies in air pollution data or incomplete modeling of the air pollution effect could be seriously exacerbated, leading to substantially attenuated effect estimates and loss of power. This attenuation will be particularly acute if, as in the 90 cities, exposure data are mostly available every six days.

Clearly, in deciding how closely to model time, we must strike a balance among competing claims. This challenge besets all methods of analysis.

NUMERICAL METHODS: DIRECT SOLUTION VERSUS BACKFITTING

In Poisson regression analysis, simultaneous estimation of a large number of time stratum effects with the regression parameters of interest may not be feasible in many software packages, which typically use quasi-Newton methods to maximize the Poisson log-likelihood. An alternative is provided by a Gauss-Siedel method—an iterative approach in which one alternates between

- fitting time stratum effects assuming that other parameters are known and equal to their current estimates, and
- fitting the remaining terms of the model assuming time stratum effects are known and equal to their current estimates.

A variation of this method is the backfitting algorithm proposed by Hastie and Tibshirani (1990) in the more general context of GAMs. In the context of the stratification (step function) model for time effects, stratum effects would be estimated within the iteratively reweighted least squares (IRWLS) fitting procedure, in step (1), by calculating stratum means of working *y* values after subtraction of the remaining fitted effects.

There are two disadvantages with the use of the Gauss-Siedel approach to maximizing the Poisson likelihood:

- This approach is known to be slow in converging when second derivatives between the two sets of parameters are appreciable (ie, when partial collinearities exist between the design matrix for time effects and design matrix for other effects).
- The full inverse second-derivative matrix is not calculated so that the standard errors of the parameters of interest are not readily available. Calculation of correct standard errors may involve large matrices. Approximate methods are available, but these may break down under strong collinearity.

These difficulties are avoided by the conditional likelihood approach because the large numbers of parameters representing time effects are eliminated from the likelihood by the conditional argument.

EXTENDING MODEL FOR TIME EFFECTS

A clearly undesirable feature of the stratification approach is that a step function is, at best, a crude approximation to the real time course of rates—particularly if time strata are wide. The methods used in the two current reports may be regarded as useful elaborations of this simple stratification approach.

Dominici, Schwartz, and their colleagues used GAMs to model both time and temperature effects. As indicated, GAMs use the Gauss-Siedel numerical method, and the computations are very similar to those outlined for simple stratification by time. At each step, however, instead of estimating time effects by calculating stratum means of working y values within disjoint time strata, local regression smoothers or smoothing splines are used.

Both reports consider the alternative of replacing the step function model with a natural cubic spline model and a limited number of knots. This remains a simple parametric Poisson regression model and has two desirable features:

- No computational problems arise in fitting all parameters of the model simultaneously using standard software.
- Inference (by maximum likelihood) is well understood (at least when no autocorrelation exists between errors) and standard errors of parameters of interest are readily available.

However, dependence of results on the choice of knots that define the spline functions is generally regarded as a less attractive feature of this approach. For modeling temperature effects, the placement of knots may be particularly influential and introduces some danger of bias (Schwartz et al; Schwartz 2003). Alternative approaches have therefore been explored. Schwartz and colleagues considered a more extensive set of alternatives including, in addition to GAMs, random effects analyses of richly parameterized spline models and a bidirectional case-crossover approach. This latter method is approximately equivalent to use of a running mean smoother within the Gauss-Siedel iteration; it is not so different from GAMs as might appear at first sight.

HOW MUCH SMOOTHING?

The problem that initially prompted this series of reanalyses was the discovery by the Johns Hopkins and other investigators (Dominici et al 2002; Ramsey et al 2003). Specifically, default convergence criteria for the Gauss-Siedel procedure built into S-Plus implementation of the GAM approach were not satisfactory for the strong collinearities typically found in this application. This means that some previously published estimates may be incorrect. In consideration of this, however, two further questions have recently received renewed attention: how much smoothing is appropriate and how standard errors of parameters of interest are estimated.

The first of these questions is more relevant to the important question of how well we might have controlled for confounding (ie, to the bias of effect estimates as estimates of causal effects). As indicated previously, modeling time by too smooth a function risks residual confounding and autocorrelation of errors. Attempting to model the time effect too closely, however, may attenuate the effect because some of the causal effect of air pollution is taken into the time effect and eventually leads to increased standard errors of air pollution effects estimated from short fluctuations. This has been the subject of substantial discussion in the time-series literature. Although the NMMAPS investigators have been thoughtful in their approach to this question, it continues to be a challenge in time-series studies.

For fitting GAMs, the degree of smoothing is chosen a priori, much as we would choose a suitable interval for stratification by time. Similarly, for parametric models using natural splines, the amount of smoothing is decided a priori by the spacing of knots. As is pointed out by Schwartz and colleagues and is further illustrated by our discussion of the stratification approach, we cannot expect to estimate an appropriate degree of smoothing from the data alone. Although the NMMAPS investigators do not in general follow this approach, it is what is attempted in many analyses of such data. In purely parametric approaches, the number of knots (and hence parameters) can be increased in a step-wise fashion and, analogously, the span of smoother within the GAM approach can be progressively shortened. Either case requires a decision on when an appropriate degree of smoothing has been achieved. Some have advocated use of the Akaike information criterion or other cross-validation methods to choose an optimal degree of smoothing. But optimality is defined in such approaches in terms of optimal prediction of a future dataset. This is different from our concern, which is with adequate control for confounding to better inform judgments about causality. Others have suggested choosing the degree of smoothing so that there is no autocorrelation of errors. Positive autocorrelation indicates *oversmoothing*, resulting in incomplete modeling of the time trajectory and, therefore, inadequate control for confounding. Thus, fitting time in too much detail may obscure some of the genuine effect of air pollution. We nevertheless believe that a complete exploration of confounding requires modeling time effects in as much detail as is possible without substantial consequences for the standard errors of parameters of interest.

Beyond controlling for time, of particular interest is the question of how fully to control for the potential effects of weather. The NMMAPS investigators and others (Mackenbach et al 1993; Samet et al 1997; Keatinge and Donaldson 2000; Ferreira Braga et al 2001; Curriero et al 2002) have explored the potential confounding effects of weather in great detail and have found them to be substantial if not controlled adequately. The current analyses continue to include terms for weather, and the sensitivity

analyses have further attempted to make estimates in the absence of extreme weather days (eg, temperatures above or below a certain point). Smaller estimates of effect result from the newly revised analyses, however, and the possibility of subtler effects of weather within the normal climatic ranges continues. Thus, further exploration of weather effects is merited (for example by considering correlated or cumulative effects of multiday temperature or humidity).

In general, GAMs are expected to give biased estimates of the regression parameters of interest. The extent of bias reduces with the roughness of the fitted curve, but with an attendant cost in terms of standard errors and, possibly, in reduction of effect due to imperfect exposure measurement. Thus, there continues to be a need to further explore the sensitivity of conclusions to this aspect of modeling.

RANDOM EFFECTS AND PENALIZED LIKELIHOOD

A recent development pointed out by Schwartz and colleagues is use of the random effects formulation for penalized spline models. Here the general idea is to specify the time model as a richly parameterized spline model. Typically, the method uses many more knots than would be used in a conventional parametric analysis but many fewer than by use of the data points themselves as knots (as is the case with the spline smoothers offered in GAM). Inferential problems associated with specifying too many nuisance parameters are avoided by introducing the parameters that correspond to the coefficients of spline basis functions as random effects in a mixed model. By use of the Laplace integral approximation, maximum likelihood estimation of fixed effects (including the parameters of interest) can be approximated by the method of maximum penalized likelihood. Because the number of additional parameters is not too large, calculation of correct standard errors is not onerous. The tradeoff, however, between bias and variance of the air pollution effect estimates with varying degrees of smoothness in GAMs applies equally to this approach.

Some would claim that the attraction of the random effects method is that the degree of smoothing is controlled by a hyperparameter that, in principle, may be estimated from the data using restricted maximum likelihood (REML). Again, however, the appropriateness of this hyperparameter to the task in hand is open to question. The information available for estimation of the hyperparameter that controls smoothness of the time effect is drawn from overdispersion and autocorrelation of residuals. While such aspects are relevant to the potential for confounding, they should not alone determine how closely we should control for time. In the current reports, Schwartz and colleagues have not followed this course but instead have

chosen the hyperparameter on a priori grounds (based on considerations of the corresponding effective degree of freedom). They have used this approach simply as an alternative to GAM that avoids the approximation of standard errors necessitated by the Gauss-Siedel iteration. Our remarks concerning choice of an appropriate degree of smoothing apply equally here.

A Bayesian argument can also be advanced for this mixed model approach. The choice of a fixed value for the smoothness hyperparameter can be taken as reliance upon a strong a priori degree of belief. The strong interdependence between the mean model for time effects and assumptions concerning autocorrelation of errors have been highlighted by Davidian and Giltinan (1995, p 133).

CASE-CROSSOVER ANALYSES

A considerable part of the report of Schwartz and colleagues is taken up with a discussion of the case-crossover approach. As indicated above, we are not convinced that this method represents a radically different approach from GAM. Indeed, if one were to implement, in GAM software, a smoother which calculates a running mean of the observations at times $t-15, t-14,\ldots,$ $t-8, t-7, t, t+7, t+8,\ldots, t+13, t+14$, then one would obtain estimates which would not differ by very much from those from the case-crossover analysis advocated by Schwartz and colleagues. It is not clear why this smoother should have better properties than the local regression or spline based smoothers available in other approaches. Further, as pointed out by other authors (Lumley and Levy 2000), overlapping control days in this approach means that the conditional likelihood is not a true likelihood and that standard errors based upon the second derivative of the conditional logistic log likelihood are likely to be incorrect.

Neither can this approach be justified by the idea of risk set sampling (Navidi and Weinhandl 2002). Even if such methods can be justified in terms of (possibly bias-corrected) estimating equations, development of methods for consistent estimation of standard errors of parameter estimates may be difficult because overlapping periods will induce correlations between contributions to the pseudoscore function. Thus, the bidirectional case-crossover approach may introduce as many problems as it solves. Although Schwartz and collaborators have presented simulation studies suggesting that such concerns are misplaced in the context of these studies, and such simulations can be useful, simulations are an imperfect substitute for sound theoretical results. The case-crossover approach in which case-control sets are nonoverlapping strata of days (Levy et al 2001) has more appeal in resolving the problem of standard error estimation (as described under the Stratification by Time section) while retaining asymptotic theory that accompanies this standard model.

criterion used to help determine the appropriateness of temporal control is the absence of positive correlation over time in the regression model errors (residual autocorrelation). Nevertheless, while positive autocorrelation does indicate when the control for time is not fine enough, a good case can be made for controlling time more finely than indicated by the disappearance of positive autocorrelation.

Alternatively, given the difficulty in predetermining optimal control, NMMAPS investigators and others used a range of smoothing and degrees of freedom to determine the extent to which these alternative approaches changed results. At the same time, as described in more detail in the Detailed Discussion of Analytic Methods, further sensitivity analyses are needed to evaluate effect estimates.

An additional important issue addressed by the revised analyses was improved estimation of standard errors. Estimation of standard errors is important not only for tests and confidence intervals in city-specific analyses, but also for later consideration of the heterogeneity of air pollution effects among cities. If standard errors of city-specific effect estimates are underestimated, tests are more likely to indicate the presence of heterogeneity.

DISCUSSION OF RESULTS

INDIVIDUAL-CITY RESULTS

The revised 90 individual-city mortality results show that, in general, the estimates of PM effect are shifted downward and the confidence intervals are widened. One consequence of this shift is that four more cities have effect estimates that are negative or zero than was the case originally.

COMBINED RESULTS ACROSS CITIES

As in the original report, a second stage metaanalysis was used to combine results on effect estimates of particulate and other pollutants on health outcomes across cities. The question here is what impact the revised analyses have had on conclusions: What is the mean estimated effect? How much heterogeneity is there across cities? What are its determinants?

Overall Mean Effect Estimate

Tightening the convergence criteria in GAM yielded a substantially lower estimate of effect of PM_{10} combined over all cities, and use of GLM with natural splines decreased the estimate further. There remained a small, but statistically significant, effect of PM_{10} at lag 1 on total mortality, now estimated to be 0.21%, with a posterior standard error of 0.06%. The Review Panel agrees with the

investigators' conclusions that the qualitative conclusions of NMMAPS II have not changed although the evidence for an effect of PM_{10} at lag 0 and lag 2 is less convincing under the new models. As reported before (Samet et al 2000), the PM_{10} effect estimates did not decrease when copollutants were introduced in the model.

A number of sensitivity analyses were undertaken in the mortality study in NMMAPS II. Results from the GLM with natural splines can be described broadly as not very sensitive to changes in modeling assumptions. For example, the degrees of freedom used for the smooth functions were changed under several different scenarios by Dominici and colleagues.[33] All the point estimates are now lower, but the pattern is similar. Making the time function smoother somewhat increased estimates, but making them less smooth or changing the smooth functions of weather did not change estimates.

For the morbidity analyses by Schwartz and colleagues, changes in methods for metaanalyses and changes in smoothing methods were potential sources of differences between new and previous combined results. This additional complexity aside, the Panel welcomed these changes. The maximum likelihood estimate of between-city variance used in the revised analysis of morbidity has better statistical properties than the previous method-of-moments estimator. The shift in emphasis from fixed effects means to those of random effects models better reflects the uncertainty in these means (although in the presence of heterogeneity, interpretation of the mean remains problematic).

Changes in the mean of 14 estimates of effect of PM_{10} on morbidity are summarized in Table 2 of the Schwartz and colleagues revised analysis. Here, results for the original analyses are compared with the results for GAMs with tighter convergence bounds (with changes in the method of combining results across cities) and for GLM with natural cubic splines. Tightening the convergence decreased the combined estimates of morbidity, and using the GLM with natural splines yielded further decreased estimates. The relative percentage decrease for the combined morbidity estimates, about 8% to 10%, is smaller than the decrease in mortality estimate reported by Dominici and colleagues. There is still evidence of a small but statistically significant estimate of PM_{10} effect on morbidity from COPD and CVD, based on the model that uses unconstrained distributed lags. (While several types of lag models were investigated in NMMAPS II, members of the original Special Panel for NMMAPS preferred unconstrained distributed lags.[34] Much of the revised analysis concentrates on this model as well.)

---

33. Dominici et al, Figure 10; Samet et al 2000, Figure A.1, p 53.

34. Samet et al 2000, p 74.

The evidence for an effect of PM_{10} on pneumonia was substantially weaker using the GLM with natural splines. Schwartz and colleagues argued fairly convincingly to the Panel that this is due to the particularly marked seasonality of pneumonia deaths. The Panel was less convinced, however, of the argument that this model is likely to have led to error for this outcome. The dependence of results for pneumonia on method of analysis in our view suggests caution in interpreting them.

The new estimates using the penalized spline method of fitting GAMs (as currently implemented in R) generally fell midway between the new GAM estimates and the estimates from GLM with natural splines. More research is needed to understand the performance of the penalized spline method with these types of data.

It is tempting to speculate that the less smooth time function in the mortality analysis (compared with the morbidity analysis) was implicated in the greater change of the mean effect estimate on reanalysis. Calculations for city-specific CVD hospitalizations used smoothing on time varying from less than 1 *df*/year to about 4 *df*/yr. The analyses also showed a clear pattern of greater change in cities where more degrees of freedom were used, but this pattern was not present for COPD or pneumonia. Neither does the smoothness of time function explain the sharply different results for pneumonia compared with COPD and CVD. Further, the mortality results for varied smoothness of time and weather functions[35] do not show an obvious pattern of change depending on smoothness. Thus, if smoothness of the time function is a determinant of impact of reanalysis, it is not predominant, and other factors need to be considered.

Heterogeneity

Each city-specific estimate provides weak information about the PM_{10} effect. If they are all measuring the same unknown quantity, averaging across cities should allow a more precise estimate of this quantity. Two ways of assessing whether the data are compatible with this hypothesis have been used in NMMAPS: (1) using a chi-square test for the presence of heterogeneity and (2) estimating heterogeneity as standard deviation of true city-specific effects. Both of these techniques allow separation of random variation between cities (expected by chance) from true variation that is over and above chance. Individual tests that are not significant or estimated standard deviations that are small favor focusing on an overall national mean of city-specific estimates. If individual tests are significant or the standard deviations are large, interpretation of the national mean becomes more problematic.

How did the reanalyses change evidence for heterogeneity? For the mortality analysis, the 90 city-specific estimates were different and usually smaller, and the estimated standard errors were generally larger using the GLM with natural splines. (This model was preferred for standard error estimation at the time of this revised analysis because correct standard error estimates using GAM were not yet available.) These results minimize the evidence for heterogeneity among the cities. Specifically, the new chi-square tests for heterogeneity are not significant, and the estimate from the Bayesian two-stage model for among-city standard deviation is lower, dropping from 0.112 (using the original GAM method) to 0.075 using GLM with natural splines. More detailed sensitivity analyses have been carried out to investigate sensitivity of the overall national mean of the PM_{10} effect estimate to prior specification of between-city variance.[36] The Panel agrees that the estimate of national mean PM_{10} effect is reassuringly robust to specification of prior, but estimated heterogeneity is probably more sensitive to this. The estimated between-city standard deviation of 0.075[37] should be taken as less certain than suggested by its credible 95% interval of 0.021 to 0.198.

The revised analysis thus shows less evidence for heterogeneity than the original analysis. This does not prove its absence, however, because power of the test to detect heterogeneity is limited (and reduced by increased standard errors in city-specific estimates) and because estimates of standard deviation are imprecise. It is thus prudent to consider interpretation in the presence of heterogeneity. Heterogeneity might be due to unmeasured confounding, in which case the combined mean is only meaningful if this confounding is assumed to be independent of pollution, an assumption that cannot be tested by the data. Alternatively, heterogeneity might be due to true variation in risk increments per unit of PM_{10} in different cities. In any case, heterogeneity complicates interpretation of the overall mean effect estimate.

The morbidity analysis yielded less information about heterogeneity because only 14 cities were investigated. Chi-square tests for heterogeneity are qualitatively similar to the NMMAPS II results:[38] some outcomes at some lags show significant heterogeneity, while others do not. As noted, the method of combining city-specific estimates more appropriately incorporated uncertainty in the overall mean effect estimate, with the result that its estimated standard errors are somewhat larger.

---

35. Dominici et al, Figure 10; Samet et al 2000, Figure A.1, p 53.

36. Dominici et al, Figure 11, Appendix.

37. Dominici et al, Table 2.

38. Schwartz et al, Table 4; Samet et al 2000, Table 35, p 38.

CONCLUSIONS AND RECOMMENDATIONS

The NMMAPS investigators have conducted a range of revised analyses, appplying alternative methods to correct shortcomings in the S-Plus GAM programming. They are to be commended for the way in which they brought these issues to light, moved expeditiously to begin to address them, and continue to conduct sensitivity analyses and explore new analytic methods.

Based on its review of their revised analyses, and those of others included in this Special Report, the Panel reached the following conclusions:

- In general, although the estimates of effect decreased substantially, the qualitative conclusions of the NMMAPS have not changed as a result of revising the analyses.

- While the alternative approaches used to model temporal effects in the revised NMMAPS analyses addressed the problems of obtaining incorrect effect estimates and standard errors when using the preprogrammed GAMs software, at this time no models can be recommended as being strongly preferred over another for use in this context.

- Findings from the revised analyses have renewed awareness of the uncertainties present in estimates of short-term air pollution effects based on time-series data.

- Although formal tests of PM effect across cities did not indicate evidence of heterogeneity because of the generally large individual-city effect standard errors, the power to assess the presence of heterogeneity was low. The possibility of heterogeneity therefore remains.

- Since the appropriate degree of control for time in these time-series analyses has not been determined, the Special Panel recommends the following:

  1. The impact of more aggressive control for time should be explored and presented as sensitivity analyses in future time-series studies (keeping in mind that the decision of how much to control for time cannot be made on purely statistical grounds).

  2. Studies to evaluate bias related to the analytic approach to smoothing and the degree of smoothing should be encouraged.

- That time effects are so strong that problems of numerical convergence can have an appreciable effect on estimates of PM effect testifies to the likelihood of important time-dependent confounders of the relation. Of such potential confounders, weather continues to be a concern. Modeling the effects of weather and its

relation to PM is a difficult problem that has been addressed by the NMMAPS investigators and others. The Panel concluded, however, that a more detailed model for the effects of weather could minimize sensitivity to the manner in which time is incorporated into the analysis and that further work on modeling weather-related factors is necessary.

IMPACT

NMMAPS

Problems with the application of GAM with default convergence criteria and the programming shortcut in standard error calculation will undoubtedly color interpretation of time-series studies in the evaluation of air pollution health effects. While estimates of effect are quantitatively smaller than those originally reported, a statistically significant overall effect estimate of PM_{10} on mortality remains. Therefore, many conclusions that were initially drawn from NMMAPS are to a large degree unchanged.

Further, because of the general increase in city-specific standard errors in the revised results, the power to detect heterogeneity among the city-specific effect estimates is reduced. The city-specific results suggest some among-city heterogeneity at face value, and four more cities have estimates of effect that are either zero or negative in the revised results. The tests provide less evidence for heterogeneity than originally, however, and suggest that heterogeneity can now be explained by chance.

AIR POLLUTION TIME-SERIES STUDIES

Compared to randomized experimental studies in which the investigator controls the intervention, findings from observational studies (such as time series) are always susceptible to uncontrolled biases and must therefore be interpreted cautiously. Observational air pollution and health studies are no exception. Uncovering the problem of using inappropriate default convergence criteria in the GAM function has again highlighted potential confounding bias in air pollution time-series studies. As in the analyses of many observational studies, avoidance of confounding bias typically entails appropriate measures of the confounding factors as terms in a regression analysis and correct specification of these factors. Determination of the appropriate degree of smoothing time in air pollution time-series studies has become a central issue. Overly aggressive smoothing may allow residual confounding, while inadequate smoothing may allow some or all of the air pollution effect to be incorporated into the smooth term. The best

method for selecting the appropriate degree of smoothing needed to control confounding bias remains to be determined. Furthermore, as is clear from the discussion of approaches to handling time, there is no gold standard. These issues introduce uncertainty into time-series studies that has motivated ongoing work to gain much needed insight. At this time, demonstration of sensitivity, or lack of it, to a range of sensible smoothing choices seems a reasonable approach.

STATISTICAL SOFTWARE

The problem in applying GAMs has sent a cautionary note to investigators using statistical software (Dominici et al 2002). Clearly, application of the S-Plus GAM function underestimated standard errors in air pollution time-series studies, and, until June 2002, the default convergence criteria provided incorrect effect estimates. The nearly ubiquitous use of GAMs in this setting reflects one of the hazards of taking a standardized approach to analysis without verifying the detailed functioning of a given application. Clearly, widespread use by applied biostatisticians and epidemiologists, as in this case, does not guarantee that the software or algorithm has no drawbacks. To their credit, investigators at Johns Hopkins continued to test their models and as a result brought the issue of the default convergence criteria to light. Looking ahead, analysts need to ensure that statistical software is appropriate for a given application or to the use of different software (Lumley and Sheppard 2003; Samet et al 2003). Again, the use of sensitivity analyses is included among these cautions (in this case addressing sensitivity, or the lack of it, in software tuning parameters and their defaults).

IMPACT CALCULATIONS

Common practice has come to use effect estimates from observational air pollution studies to calculate the impact of air pollution on a large population such as that of the United States. If effect estimates from the NMMAPS 90 cities mortality study were applied, the revised impact would be approximately half of the estimated impact derived using the original effect estimates. This example reinforces the need to qualify estimates of impact by emphasizing the assumptions and uncertainties on which the estimates are based.

LONG-TERM EFFECTS STUDIES

Some have noted that the calculated health impact of short-term air pollution concentrations based on time-series studies is substantially smaller than the calculated impact of long-term air pollution concentrations based on cohort studies (Pope et al 2002; Zeger et al, 2003). Because of the vastly larger number of time-series studies performed, however, assessors of risk from air pollution have often been more confident in their findings than in findings from the few cohort studies. Because the problem with applying GAMs has involved primarily the time-series studies and correction of the problem has generally decreased estimates of effect from these studies, more emphasis on cohort studies can be expected. Further, uncertainty regarding the estimates of effect from time-series studies can also be expected to place additional emphasis on long-term air pollution studies, on studies of natural experiments (the so-called quasiexperimental studies), and on human experimental and animal toxicologic studies.

REFERENCES

Curriero FC, Heiner KS, Samet JM, Zeger SL, Strug L, Patz JA. 2002. Temperature and mortality in 11 cities of the eastern United States. Am J Epidemiol 155:80–87.

Davidian M, Giltinan DM. 1995. Nonlinear Models for Repeated Measurement Data. Chapman and Hall, London, England.

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On generalized additive models in time series studies of air pollution and health. Am J Epidemiol 156:193–203.

DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. Control Clin Trials 7(3):177–188.

Ferreira Braga AL, Zanobetti A, Schwartz J. 2001. The time course of weather-related deaths. Epidemiology 12(6):662–667.Hastie TJ, Tibshirani RJ. 1990. Generalized Additive Models. Monographs on Statistics and Applied Probability 43. Chapman and Hall, New York NY.

Keatinge WR, Donaldson GC. 2000. Mortality related to cold and air pollution in London after allowance for effects of associated weather patterns. Environ Res 86:209-216.

Levy D, Lumley T, Sheppard L, Kaufman J, Checkoway H. 2001. Referent selection in case-crossover analyses of acute health effects of air pollution. Epidemiology 12:186–192.

Lumley T, Levy D. 2000. Bias in the case-crossover design: Implications for studies of air pollution. Environmetrics 11:689–704.

Lumley T, Sheppard L. 2003. Time series analyses of air pollution and health: Straining at gnats and swallowing camels? Epidemiology 14:11–12.

Mackenbach JP, Looman CW, Kunst AE. 1993. Air pollution, lagged effects of temperature, and mortality: The

Netherlands 1979-87. J Epidemiol Community Health 47:121–126.

Navidi W, Weinhandl E. 2002. Risk set sampling for case-crossover designs. Epidemiology 13:100–105.

Pope CA III, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 287:1132–1144.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity on generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Samet JM, Dominici F, McDermott A, Zeger SL. 2003. New problems for an old design: Time series analyses of air pollution and health. Epidemiology 14:11–12.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwartz J, Zanobetti A. 2000. The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Samet JM, Zeger SL, Kelsall JE, Xu J, Kalkstein LS. 1997. Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Air Pollutants. The Phase I.B Report of the Particle Epidemiology Evaluation Project. Special Report. Health Effects Institute, Cambridge MA.

Zeger SL, Dominici F, McDermott A, Samet JM. 2003. Modeling the health effects of environmental exposures: Particulate air pollution and mortality in urban America. In: Monitoring the Health of Populations: Statistical Methods for Public Health Surveillance (R Brookmeyer and D Stroup, eds). Oxford University Press, New York NY. In press.

ABBREVIATIONS AND OTHER TERMS

CI	confidence interval
COPD	chronic obstructive pulmonary disease
CVD	cardiovascular disease
df	degrees of freedom
GAM	generalized additive model
GLM	generalized linear model
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
PM	particulate matter
PM_{10}	PM less than 10 μm in aerodynamic diameter

Revised Analyses of Selected Time-Series Studies

Revised Analyses of the National Morbidity, Mortality, and Air Pollution Study, Part II

Continued

Commentary on Revised Analyses of Selected Studies

A Special Panel of the Health Review Committee

List of Original Publications of Selected Time-Series Studies

Original Publication by First Author	Location in This Report
Atkinson RW, Anderson HR, Sunyer J, Ayres J, Baccini M, Vonk JM, Boumghar A, Forastiere F, Forsberg B, Touloumi G, Schwartz J, Katsouyanni K. 2001. Acute effects of particulate air pollution on respiratory admissions. Results from APHEA 2 Project. Am J Respir Crit Care Med 164:1860–1866.	Atkinson p 81
Braga ALF, Zanobetti A, Schwartz J. 2001. The lag structure between particulate air pollution and respiratory and cardiovascular deaths in 10 US cities. J Occup Environ Med 43:927–933.	Schwartz p 211
Burnett RT, Brook J, Dan T, Delocla C, Philips O, Cakmak S, Vincent R, Goldberg MS, Krewski D. 2000. Association between particulate- and gas-phase components of urban air pollution and daily mortality in eight Canadian cities. Inhalation Toxicol 12(Suppl 4):15–39.	Burnett p 85
Daniels MJ, Dominici F, Samet JM, Zeger SL. 2000. Estimating particulate matter-mortality dose-response curves and threshold levels: An analysis of daily time-series for the 20 largest US cities. Am J Epidemiol 152:397–406.	Dominici p 91
Dominici F, Daniels M, Zeger SL, Samet JM. 2002. Air pollution and mortality: Estimating regional and national dose-response relationships. J Am Stat Assoc 97:100–111.	Dominici p 91
Fairley D. 1999. Daily mortality and air pollution in Santa Clara County, California: 1989–1996. Environ Health Perspect 107:637–641.	Fairley p 97
Gold DR, Litonjua A, Schwartz J, Lovett E, Larson A, Nearing B, Allen G, Verrier M, Cherry R, Verrier R. 2000. Ambient pollution and heart rate variability. Circulation 101:1267–1273.	Gold p 107
Goldberg MS, Bailar JC III, Burnett RT, Brook JR, Tamblyn R, Bonvalot Y, Ernst P, Flegel KM, Singh RK, Valois M-F. 2000. Identifying Subgroups of the General Population That May Be Susceptible to Short-Term Increases in Particulate Air Pollution: A Time-Series Study in Montreal, Quebec. Research Report 97. Health Effects Institute, Cambridge MA.	Goldberg p 113
Goldberg MS, Burnett RT, Bailar JC III, Brook J, Bonvalot Y, Tamblyn R, Singh R, Valois M-F. 2001. The association between daily mortality and ambient air particle pollution in Montreal, Quebec. 1. Nonaccidental mortality. Environ Res 86:12–25.	Goldberg p 113
Goldberg MS, Burnett RT, Bailar JC III, Brook J, Bonvalot Y, Tamblyn R, Singh R, Valois M-F, Vincent R. 2001. The association between daily mortality and ambient air particle pollution in Montreal, Quebec. 2. Cause-specific mortality. Environ Res 86:26–36.	Goldberg p 113
Goldberg MS, Burnett RT, Bailar JC III, Tamblyn R, Ernst P, Flegel K, Brook J, Bonvalot Y, Singh R, Valois M-F, Vincent R. 2001. Identification of persons with cardiorespiratory conditions who are at risk of dying from the acute effects of ambient air particles. Environ Health Perspect 109(Suppl 4):487–494.	Goldberg p 113
Goldberg MS, Burnett RT, Brook J, Bailar JC III, Valois M-F, Vincent R. 2001. Associations between daily cause-specific mortality and concentrations of ground-level ozone in Montreal, Quebec. Am J Epidemiol 154:817–826.	Goldberg p 113
Hoek G, Brunekreef B, Fischer P, van Wijnen J. 2001. The association between air pollution and heart failure, arrhythmia, embolism, thrombosis, and other cardiovascular causes of death in a time series study. Epidemiology 12:355–357.	Hoek p 133
Hoek G, Brunekreef B, Verhoeff A, van Wijnen J, Fischer P. 2000. Daily mortality and air pollution in The Netherlands. J Air Waste Manage Assoc 50:1380–1389.	Hoek p 133

Continued

Original Publication by First Author (*Continued*)   Location in This Report

Janssen NAH, Schwartz J, Zanobetti A, Suh HH. 2002. Air conditioning and source-specific particles as modifiers of the effect of PM_{10} on hospital admissions for heart and lung disease. Environ Health Perspect 110:43–49.

Zanobetti p 241

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiology 12:521–531.

Katsouyanni p 157

Klemm RJ, Mason RM Jr, Heilig CM, Neas LM, Dockery DW. 2000. Is daily mortality associated specifically with fine particles? Data reconstruction and replication of analyses. J Air Waste Manage Assoc 50:1215–1222.

Klemm p 165

Laden F, Neas LM, Dockery DW, Schwartz J. 2000. Association of fine particulate matter from different sources with daily mortality in six US cities. Environ Health Perspect 108:941–947.

Schwartz p 219

Le Tertre A, Medina S, Samoli E, Forsberg B, Michelozzi P, Boumghar A, Vonk JM, Bellini A, Atkinson R, Ayres JG, Sunyer J, Schwartz J, Katsouyanni K. 2002. Short-term effects of particulate air pollution on cardiovascular diseases in eight European cities. J Epidemiol Community Health 56:773–779.

Le Tertre p 173

Lippmann M, Ito K, Nádas A, Burnett RT. 2000. Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations. Research Report 95. Health Effects Institute, Cambridge MA.

Ito p 143

Mar TF, Norris GA, Koenig JQ, Larson TV. 2000. Associations between air pollution and mortality in Phoenix, 1995–1997. Environ Health Perspect 108:347–353.

Mar p 177

Moolgavkar SH. 2000. Air pollution and daily mortality in three US counties. Environ Health Perspect 108:777–784.

Moolgavkar p 183

Moolgavkar SH. 2000. Air pollution and hospital admissions for chronic obstructive pulmonary disease in three metropolitan areas in the United States. Inhalation Toxicol 12(Suppl 4):75–90.

Moolgavkar p 183

Moolgavkar SH. 2000. Air pollution and hospital admissions for diseases of the circulatory system in three US metropolitan areas. J Air Waste Manage Assoc 50:1199–1206.

Moolgavkar p 183

Ostro BD, Broadwin R, Lipsett MJ. 2000. Coarse and fine particles and daily mortality in the Coachella Valley, California: A follow-up study. J Expo Anal Environ Epidemiol 10:412–419.

Ostro p 199

Samet JM, Dominici F, Zeger SL, Schwartz J, Dockery DW. 2000. The National Morbidity, Mortality, and Air Pollution Study, Part I: Methods and Methodologic Issues. Research Report 94. Health Effects Institute, Cambridge MA.

Dominici p 91,
 Schwartz p 219

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwartz J, Zanobetti A. 2000. The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Dominici p 9,
 Schwartz pp 25 & 211,
 Zanobetti p 241

Samoli E, Schwartz J, Wojtyniak B, Touloumi G, Spix C, Balducci F, Medina S, Rossi G, Sunyer J, Bacharova L, Anderson HR, Katsouyanni K. 2001. Investigating regional differences in short-term effects of air pollution on daily mortality in the APHEA project: A sensitivity analysis for controlling long-term trends and seasonality. Environ Health Perspect 109:349–353.

Samoli p 205

Schwartz J, Dockery DW, Neas LM. 1996. Is daily mortality associated specifically with fine particles? J Air Waste Manage Assoc 46:927–939.

Schwartz p 219

Schwartz J. 2000. Assessing confounding, effect modification, and thresholds in the association between ambient particles and daily deaths. Environ Health Perspect 108:563–568.

Schwartz p 211

Schwartz J. 2000. Harvesting and long term exposure effects in the relation between air pollution and mortality. Am J Epidemiol 151:440–448.

Schwartz p 219

Continued

Original Publication by First Author (*Continued*)	Location in This Report
Schwartz J. 2000. The distributed lag between air pollution and daily deaths. Epidemiology 11:320–326.	Schwartz p 211
Sheppard L, Levy D, Norris G, Larson TV, Koenig JQ. 1999. Effects of ambient air pollution on nonelderly asthma hospital admissions in Seattle, Washington, 1987–1994. Epidemiology 10:23–30.	Sheppard p 227
Wichmann H-E, Spix C, Tuch T, Wölke G, Peters A, Heinrich J, Kreyling WG, Heyder J. 2000. Daily Mortality and Fine and Ultrafine Particles in Erfurt, Germany. Part I: Role of Particle Number and Particle Mass. Research Report 98. Health Effects Institute, Cambridge MA.	Stölzel p 231
Zanobetti A, Schwartz J, Dockery DW. 2000. Airborne particles are a risk factor for hospital admissions for heart and lung disease. Environ Health Perspect 108:1071–1077.	Zanobetti p 241
Zanobetti A, Schwartz J, Samoli E, Gryparis A, Touloumi G, Atkinson R, Le Tertre A, Bobros J, Celko M, Goren A, Forsberg B, Michelozzi P, Rabczenko D, Ruiz EA, Katsouyanni K. 2002. The temporal pattern of mortality responses to air pollution: A multicity assessment of mortality displacement. Epidemiology 13:87–93.	Zanobetti p 249
Zeger SL, Dominici F, Samet J. 1999. Harvesting-resistant estimates of air pollution effects on mortality. Epidemiology 10:171–175.	Dominici p 91

Acute Effects of Particulate Air Pollution on Respiratory Admissions

Richard W Atkinson, H Ross Anderson, Jordi Sunyer, Jon Ayres, Michela Baccini,
Judith M Vonk, Azzedine Boumghar, Francesco Forastiere, Bertil Forsberg,
Giota Touloumi, Joel Schwartz, and Klea Katsouyanni

ABSTRACT

The Air Pollution and Health: A European Approach 2 (APHEA2*) project investigated the short-term effects of air pollution on daily mortality and hospital admissions in 29 European cities. Particulate matter less than 10 μm in diameter (PM_{10}) and black smoke (BS) were investigated, and the results have been published for all-cause mortality and both respiratory and cardiovascular hospital admissions. The original analyses, using locally weighted smoothers (LOESS) within a generalized additive model (GAM) framework, were conducted using S-Plus software with default convergence criteria settings. The sensitivity of the original respiratory admissions results to more stringent convergence criteria and alternative smoothing functions (natural splines with the same degree of freedom) have since been investigated. For both PM_{10} and BS, the results for hospital admissions for asthma in children and adults, for chronic obstructive pulmonary disease (COPD), and for all respiratory causes in persons over the age of 65 were insensitive to the specification of the convergence criteria. Further, they were generally unaffected by the choice of smoothing function used in GAM.

INTRODUCTION

The APHEA2 project investigated the short-term effects of air pollution on daily mortality and hospital admissions in 29 European cities. The project aimed to study the consistency of air pollution effect estimates between cities and, where significant heterogeneity between estimates was found, to investigate whether this variability could be explained by city-level factors such as population structure, environmental conditions, and health of the

population. Results for particles and daily mortality and hospital admissions have been published [Katsouyanni et al 2001; Le Tertre et al 2002; Atkinson et al 2001].

This short communication report has been prepared at the request of the US Environmental Protection Agency and the HEI Revised Analyses Review Panel. The request followed concerns that default convergence parameters commonly used in statistical software for GAM analyses were not stringent enough. This report presents results of a reanalysis of the particle and respiratory admissions data [Atkinson et al 2001] taking into account the new recommendations for convergence criteria. Results for PM_{10} and BS are presented here. A reanalysis of the data using natural cubic splines is also discussed.

METHODS

The statistical analyses followed a protocol devised by the APHEA2 project members. For each outcome studied, a GAM was constructed to model the daily number of admissions as a function of seasonal patterns, meteorological factors, and other explanatory variables. LOESS were used to model the seasonal patterns found in the outcome series and the nonlinear dependency of admissions on temperature and humidity. Time series of respiratory admissions can display episodic patterns, suggesting that the amount of smoothing required varies over the duration of the time series. This adjustment was accomplished by using smoothing terms for specific time periods within the overall series. Models were fitted using quasi-likelihood assuming constant overdispersion over time, which allowed a relaxation in assumptions for distribution of the model residuals. The overdispersion parameter was estimated using Pearson residual χ^2.

Serial correlation is often present in respiratory admission time series; it is due to external confounding factors such as seasonal patterns, weather, and respiratory epidemics rather than an inherent property of the series. The objective of modeling was therefore to account for the confounding factors responsible for inducing this serial

---

correlation without also removing the short-term effects of air pollution on the outcome under study. The principal tool used to assess this objective was the sample partial autocorrelation plot. Respiratory admissions series have strong serial correlation in the first few lags, so the objective was to reduce the partial autocorrelation estimates beyond the first few lags to nonsignificant random variations about zero. Adjustments for the remaining serial correlation were made using an autoregressive error model. Assessment of the suitability of the model was supplemented with time-series plots of fitted values and Pearson residuals to ensure that no strong seasonal patterns remained in the model residuals.

After controlling for seasonal and long-term trends in the time series, weather terms were incorporated into the model. The exact functional form of both temperature and humidity as well as the lags of weather variables were chosen using the Akaike information criterion (AIC) and diagnostic plots. Generally, same-day and lagged temperature values were included, using either parametric or nonparametric forms, according to the maximum reduction in AIC achieved. The appropriate functional form of the influenza variable(s) was also added to the model. Indicator variables for the day of the week, public holidays, school holidays (if appropriate), and any unusual events were also included in the models.

This modeling procedure was carried out for each of the outcome series (four disease/age groups and eight cities). As a result, the core models for each time series were not necessarily the same. The final step in the modeling procedure was the addition of the pollutant, at which stage the model diagnostics were rechecked. It has been shown that log-transformed air pollution measures can give the best fit in settings in which high levels of air pollution are present ($> 150\ \mu g/m^3$ for particles). Because one aim of the study was to investigate heterogeneity, only linear terms were used for the pollutant measures. Thus the analyses were restricted to days with particle levels of less than 150 $\mu g/m^3$. For Milan and Rome, where only total suspended particles (TSP) measures were available, the days for analysis were selected using rescaled TSP data (conversion factor 0.75). PM_{13} measures available in Paris were considered to be close to PM_{10} measures.

The original analyses were conducted using S-Plus software using the default convergence criteria. The revised analysis used more stringent criteria (Table 1). In

Table 1. Parameters Controlling Convergence: Default and Revised, More Stringent Values

Parameter	Default	More Stringent
ε	10^{-3}	10^{-14}
M	10	1000
ε_{bf}	10^{-3}	10^{-14}
M_{bf}	10	1000

all other respects the models were identical to those used in the original analyses. To further test the sensitivity of the results to the method of analysis, the models for PM_{10} and BS were run again using natural cubic splines (the ns function in S-Plus). The degrees of freedom used in the splines were similar to those used for LOESS in the original models and the more stringent convergence criteria were also retained.

RESULTS

Tables 2 and 3 present the effects of changing the convergence criteria used in the S-Plus software from the default values to the more stringent values. For each outcome in each city, the original and revised estimates are given for 10 $\mu g/m^3$ increments in PM_{10} or BS. The particle effect estimates were largely unaffected by the more stringent criteria used in calculation of the regression estimates and their standard errors. This conclusion applies to each of the four categories of disease and age studied. Results for BS showed a similar pattern, excepting that the estimates for asthma admissions in children 0 to 14 years of age were slightly more variable than those of the other groups.

The sensitivity of results for PM_{10} to the smoothing function used (together with the more stringent convergence criteria) is illustrated in Figure 1. Each graph presents the results from the original analysis and from the revised analyses using the more stringent convergence criteria and using natural splines rather than LOESS. The figures clearly show that the original results for PM_{10} are mostly insensitive to both the convergence criteria and the smoothing function used. The results for BS are similar (data not shown).

Table 2. Comparison of Results for PM_{10} Using Default and Revised Convergence Criteria

Outcome/City or Country	Model	%[a]	95% CL[b]		%[a]	95% CL[b]		%[a]	95% CL[b]		%[a]	95% CL[b]	
		Asthma						COPD + Asthma			All Respiratory Diseases		
		Ages 0–14 years			Ages 15–64 years			Ages 65+ years			Ages 65+ years		
Barcelona	Original	2.7	−4.9	10.9	0.4	−3.5	4.4	2.6	1.0	4.3	2.0	0.8	3.1
	Revised	2.7	−4.9	10.9	0.4	−3.5	4.4	2.6	1.0	4.3	1.9	0.8	3.1
Birmingham	Original	2.8	0.8	4.8	2.5	0.1	4.9	0.5	−1.4	2.6	0.9	−0.3	2.2
	Revised	3.0	1.0	5.0	2.5	0.1	4.9	0.6	−1.4	2.6	0.9	−0.3	2.2
London	Original	0.6	−0.8	2.0	1.4	−0.1	3.0	0.3	−0.8	1.5	0.4	−0.3	1.2
	Revised	0.8	−0.6	2.2	1.3	−0.2	2.8	0.3	−0.8	1.5	0.4	−0.4	1.2
Milan	Original	4.1	1.7	6.4	0.5	−2.1	3.1	1.2	0.0	2.3	1.0	0.4	1.7
	Revised	4.1	1.8	6.5	0.5	−2.1	3.1	1.2	0.0	2.3	1.0	0.3	1.7
Netherlands	Original	−0.9	−2.1	0.4	0.4	−0.9	1.8	1.1	0.5	1.7	1.2	0.7	1.6
	Revised	−0.7	−2.0	0.5	0.4	−0.9	1.7	1.1	0.5	1.7	1.2	0.7	1.6
Paris	Original	0.7	−1.5	3.0	1.2	−0.7	3.2	−0.6	−2.5	1.3	−0.1	−1.3	1.0
	Revised	0.9	−1.4	3.2	1.2	−0.8	3.2	−0.5	−2.4	1.4	−0.2	−1.3	1.0
Rome	Original	1.4	−3.2	6.1	1.4	−3.0	6.0	0.7	−1.1	2.5	0.8	−0.5	2.1
	Revised	1.4	−3.2	6.1	1.3	−3.0	5.9	0.7	−1.1	2.5	0.8	−0.5	2.1
Stockholm	Original	1.7	−6.0	10.2	5.4	−4.0	15.7	2.7	−1.5	7.1	1.7	−1.2	4.7
	Revised	1.7	−6.0	10.2	5.4	−4.0	15.7	2.7	−1.5	7.1	1.7	−1.2	4.7
Summary estimates	Original	1.2	0.2	2.3	1.1	0.3	1.8	1.0	0.4	1.5	0.9	0.6	1.3
	Revised	1.5	0.1	2.8	1.0	0.3	1.8	1.0	0.6	1.4	1.0	0.7	1.3

[a] Percentage change in mean number of admissions associated with increments in PM_{10} of 10 μg/m^3.

[b] 95% confidence limits.

Table 3. Comparison of Results for Black Smoke Using Default and Revised Convergence Criteria

Outcome/City or Country	Model	%[a]	95% CL[b]		%[a]	95% CL[b]		%[a]	95% CL[b]		%[a]	95% CL[b]	
		Asthma									All Respiratory Diseases		
		Ages 0–14 years			Ages 15–64 years			Ages 65+ years			Ages 65+ years		
Barcelona	Original	10.4	0.4	21.4	2.1	−3.0	7.5	−2.1	−4.3	0.0	−0.7	−2.3	0.9
	Revised	10.4	0.3	21.4	2.0	−3.2	7.4	−2.2	−4.3	0.0	−0.7	−2.3	0.9
Birmingham	Original	2.0	−1.9	6.0	2.8	−1.9	7.7	2.2	−1.7	6.2	2.9	0.6	5.4
	Revised	2.5	−1.4	6.6	2.7	−1.9	7.6	2.0	−1.9	6.0	3.0	0.6	5.4
London	Original	1.1	−1.3	3.6	1.8	−0.9	4.5	0.4	−1.6	2.5	−1.1	−2.4	0.3
	Revised	1.2	−1.2	3.8	1.6	−1.0	4.4	0.5	−1.5	2.6	−1.0	−2.3	0.3
Netherlands	Original	1.4	−0.4	3.3	−0.4	−2.2	1.5	0.7	−0.2	1.6	0.0	−0.7	0.7
	Revised	1.8	0.0	3.6	−0.4	−2.2	1.5	0.7	−0.2	1.6	−0.1	−0.7	0.6
Paris	Original	0.9	−0.8	2.7	0.8	−0.7	2.3	0.2	−1.3	1.6	0.5	−0.4	1.4
	Revised	1.1	−0.6	2.8	0.8	−0.7	2.4	0.2	−1.2	1.7	0.4	−0.4	1.3
Summary estimates	Original	1.3	0.3	2.4	0.7	−0.3	1.8	0.2	−0.7	1.1	0.1	−0.7	0.9
	Revised	1.6	0.5	2.6	0.7	−0.3	1.8	0.4	−0.3	1.0	0.1	−0.7	0.9

[a] Percentage change in mean number of admissions associated with increments in PM_{10} of 10 μg/m^3.

[b] 95% confidence limits.
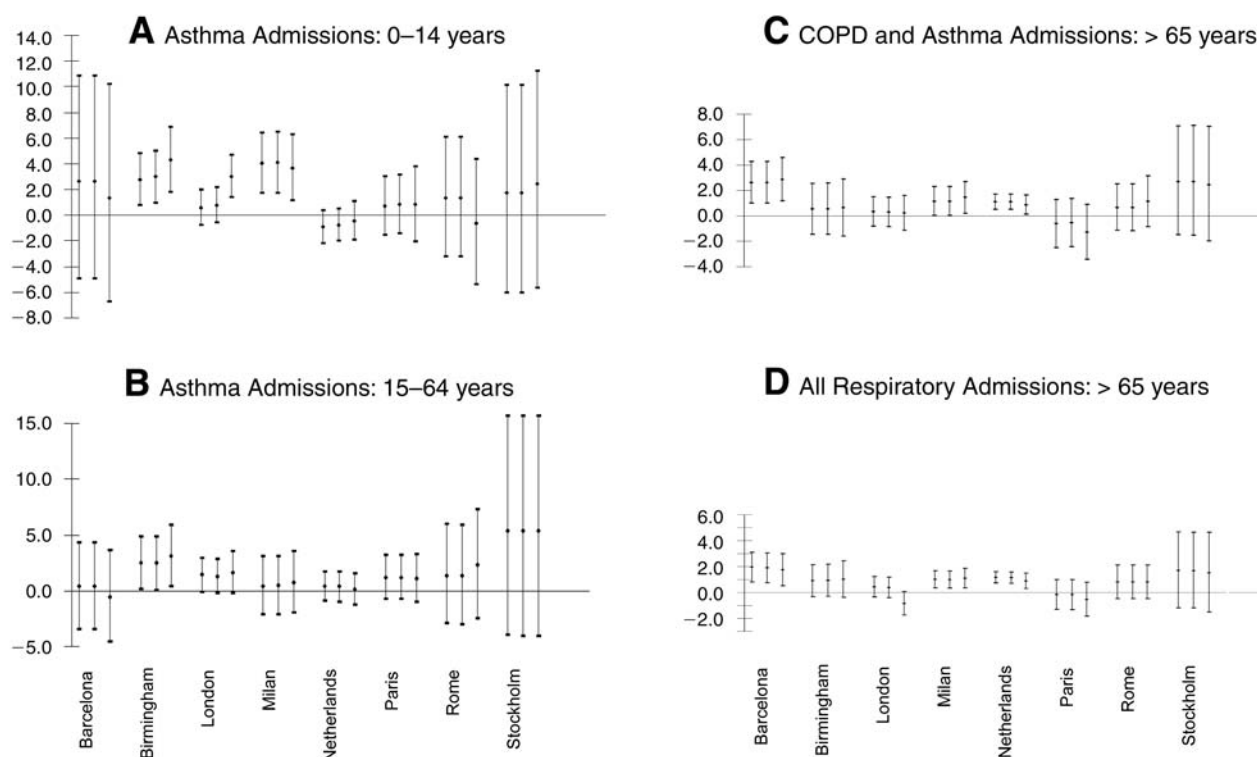
Figure 1. Sensitivity of results for PM$_{10}$ for each city. From left to right, results from the original analyses are followed by the results using the more stringent convergence criteria and then by the results using the more stringent convergence criteria together with the natural spline smoothers rather than LOESS functions. Each result given represents the percentage change (± 95% CI) as noted on y axes in the mean number of admissions associated with an increase in PM$_{10}$ of 10 µg/m^3.

DISCUSSION

These analyses show that, in the case of hospital admissions for respiratory diseases in European cities studied by the APHEA2 project, the models are robust to the convergence criteria used by the S-Plus software. The estimates were not altered significantly when different criteria were used and therefore the overall conclusions from the original study [Atkinson et al 2001] remain unaltered. Furthermore, for these time series, the results for PM$_{10}$ and BS were not particularly sensitive to the smoothing function (LOESS or natural cubic splines) used in the analysis.

REFERENCES*

Atkinson RW, Anderson HR, Sunyer J, Ayres J, Baccini M, Vonk JM, Boumghar A, Forastiere F, Forsberg B, Touloumi G, Schwartz J, Katsouyanni K. 2001. Acute effects of particulate air pollution on respiratory admissions.

Results of the APHEA2 Project. Am J Respir Crit Care Med 164:1860–1866.

Katsouyanni K et al. 2001. Confounding and Effect Modification in the Short-Term Effects of Ambient Particles on Total Mortality: Results from 29 European Cities within the APHEA 2 Project. Epidemiology 12:521–531.

Le Tertre et al. 2002. Short-Term Effects of Particulate Air Pollution on Cardiovascular Diseases in Eight European Cities. J Epidemiol Community Health 56:773–779.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
APHEA2	Air Pollution and Health: A European Approach 2
BS	black smoke
GAM	generalized additive model
LOESS	locally weighted smoothers
PM$_{10}$	particulate matter < 10 µm in diameter
TSP	total suspended particles

* Bold type identifies publication containing the original analyses revised in this short communicatin report.

Size-Fractionated Particulate Mass and Daily Mortality in Eight Canadian Cities

Richard T Burnett and Mark S Goldberg

ABSTRACT

A revised analysis of the association between size-fractionated particulate mass and nonaccidental mortality over the 11-year period 1986 to 1996 in eight Canadian cities is reported. The original publication results (Burnett et al 2000) using nonparametric, locally weighted smoothers (LOESS*) for time and weather variables using the S-Plus default convergence criteria were replaced with natural spline models. Eight smoothing selection strategies for time were purposed. Particulate associations with mortality were highly sensitive to the type of smoother and the amount of smoothing. LOESS gave higher effect estimates than natural splines. The association between coarse particulate mass and mortality was more sensitive to the amount of temporal smoothing than the fine particle–mortality association. Positive estimates of heterogeneity of particulate effect across cities were observed using LOESS while negative estimates of heterogeneity were obtained using natural splines. We conclude that strategies need to be developed for selecting the appropriate amount of smoothing of time in the conduct of time-series health studies.

INTRODUCTION

This short communication describes the results of the revised analysis of the association between size-fractionated mass of particulate matter ($PM_{2.5}$, $PM_{10-2.5}$, and PM_{10}) and nonaccidental mortality in eight Canadian cities (Montreal, Ottawa, Toronto, Windsor, Winnipeg, Edmonton, Calgary, and Vancouver) from January 1, 1986 to December 31, 1996. The paper describing the original analysis was published previously (Burnett et al 2000).

---

* A list of abbreviations and other terms appears at the end of the section.

This short communication report is part of an HEI Special Report, which also includes 20 other reports, a section on NMMAPS II, two HEI Commentaries, and an HEI Statement. Please address correspondence about this section to Dr Richard T Burnett, Biostatistics & Epidemiology Division, Health Canada, 4th floor, 2720 Riverside Drive, Ottawa ON, K1A 0K9 Canada.

The particulate and mortality data are identical to the original publication.

ORIGINAL STUDY METHODS AND RESULTS

The majority of results in the original publication used a time-series statistical approach to linking daily variations in mortality with daily variations in air pollution by removing temporal trends in the mortality, particulate and weather data prior to linking these temporally filtered variables together. This approach removes the potentially confounding effects of time from each of the time series prior to examining the effect of air pollution on mortality. We reported a 1.9% ($t = 2.8$) change in daily mortality for a 25.9 μg/m³ change in PM_{10}, 1.6% ($t = 3.1$) for a 13.3 μg/m³ change in $PM_{2.5}$, and 0.9% ($t = 1.4$) for a 12.6 μg/m³ change in $PM_{10-2.5}$ in Table 8 of the original publication. These effects are based on particulate measurements recorded on the day prior to death. The model for mortality contained a LOESS smooth of day of study with a 90-day span that was common to all eight cities. This span was selected to maximize the minimum type I error rate or *P* value of the Bartlett test (Priestly 1981) for white noise residuals among the eight cities. This approach was selected in order to ensure as best as possible that we were controlling for temporal trends in the mortality time series in each city yet allowing a common span among all cites.

All days of observation were included in the mortality–time model despite the fact that most of the particulate measurements were recorded on a six-day sampling schedule. Time trends in particles and weather were also removed from these series using a 90-day span LOESS. The temporally filtered residuals from the mortality, air pollution and weather time series were then linked using a linear regression model combining information across all cities. Daily average temperature and maximum change in barometric pressure within a day recorded on the day of death were selected as the best weather predictors based on the Akaike information criterion (AIC) coupled with a

---

stepwise regression procedure that included daily average temperature, daily average relative humidity, and barometric pressure lagged 0 and 1 days. A common weather model was selected for each city. Here, the potentially nonlinear associations between weather and mortality were described by LOESS with a 0.50 span.

A secondary analysis was also performed using a model approach that is more common in the current literature. In this approach, time, weather, and air pollution were modeled together, thus allowing the temporal cycles in each time series to compete in predicting daily mortality. Air pollution estimates based on this coadjustment approach permit information from frequencies lower than day to day, such as subseasonal or seasonal, to be included in the analysis. The results from this coadjustment analysis approach were 2.1% ($t = 3.7$) for a 25.9 µg/m^3 change in PM$_{10}$, 1.9% ($t = 4.2$) for a 13.3 µg/m^3 change in PM$_{2.5}$, and 1.2% ($t = 2.3$) for a 12.6 µg/m^3 change in PM$_{10–2.5}$.

Particulate effect estimates based on the prefiltering approach were smaller than those estimates based on the coadjustment approach, suggesting that temporal cycles in the mortality and air pollution data were positively contributing to the risk estimate.

REVISED ANALYSIS METHODS

Revised analyses of these data will be conducted in two phases. All revised analyses used the coadjustment approach because this approach is in current favor and thus our results will be comparable with other researchers. In the first phase, the influence on particulate effect estimates of adding indicator variables for day of the week was examined. Mortality rates vary little by day of the week, but there are fewer deaths on Sundays. Most other researchers now control for day of the week effects.

The influence on effect estimates of the convergence criteria used in S-Plus generalized additive model (GAM) routines was also examined. The default criteria consist of four components: a relative difference in parametric parameter estimates between local scoring iterations of 10^{-3}; a difference in the residual sum of squares between backfitting iterations of 10^{-3}; maximum iterations for the local scoring algorithm of 10; and maximum iterations for the backfitting algorithm of 10. These criteria are denoted by (10^{-3}, 10^{-3}, 10, 10). Recognizing concerns that this set of convergence criteria may not always be strict enough to find the maximum likelihood estimate, however, Dominici and colleagues (2002) suggested much stricter criteria (10^{-15}, 10^{-15}, 1000, 1000) that was used in this revised analysis.

The first phase of the revised analysis consists of determining the pooled estimate of the air pollution effect among the eight cities assuming that time (t) and weather variables were modeled by LOESS functions within the GAM framework. The model for the expected number of deaths at time t in the jth city, y_{tj}, has the form

$$E\left\{y_{tj}\right\} = \exp(lo(t, 90/4018) \\ + lo(temp_{tj}, 0.5) \\ + lo(pr_{tj}, 0.5) \\ + \beta_j AP_{tj}) \qquad , \tag{1}$$

where lo denotes the LOESS function and $temp$, pr, and AP denote daily average temperature, change in barometric pressure, and air pollution, respectively. The quantity $90/4018$ defines the span of LOESS, which is given by 90 days divided by the length of the time series (4018 days). The air pollution effect estimate for the jth of 8 cities, $\hat{\beta}_j$, is assumed to be normally distributed with expectation β and variance $\theta + \hat{v}_j$, where $\hat{v}_j$ is the estimation variance of $\hat{\beta}_j$ and θ is the variance of the true air pollution effect between cities. The estimate of the pooled air pollution effect among cities is given by

$$\hat{\beta} = \left(\sum_{j=1}^{8}(\hat{\theta} + \hat{v}_j)^{-1}\right)^{-1} \times \sum_{j=1}^{8}(\hat{\theta} = \hat{v}_j)^{-1}\beta_j,$$

where $\hat{\theta}$ is the maximum likelihood estimate of θ. The estimate of the standard error of $\hat{\beta}$ is given by

$$\left(\sum_{j=1}^{8}(\hat{\theta} + \hat{v}_j)^{-1}\right)^{-1/2}.$$

Four models for each of PM$_{2.5}$, PM$_{10–2.5}$, and PM$_{10}$ were considered, resulting in 12 model runs: default and stricter convergence criteria; with and without day-of-the-week (DOW) indicator variables.

In the second phase of the revised analysis, natural spline models were used to capture the potentially nonlinear effects of time and weather on mortality in place of the LOESS. The natural spline models are defined by the degree of freedom, which is given by the number of knots plus two. The knots are placed evenly throughout the distribution of the variable modeled. Natural spline models have been shown to perform well in time-series mortality studies in terms of bias in the air pollution effect estimate and the corresponding standard errors compared to LOESS and nonparametric spline smoothers (Dominici et al 2002; Ramsay et al 2003).

It is difficult to define an equivalent amount of smoothing produced by the natural spline compared to LOESS. Both have a concept of degree of freedom, but nonparametric smoothers are defined in a manner that fit the data better than parametric functions like natural splines for the same degree of freedom. The LOESS function was therefore not replaced with a natural spline with the same degree of freedom. Instead, the same modeling approach to select the degree of freedom for the natural splines was used as the LOESS. Namely, maximize the minimum P value of the Bartlett test for white noise among the eight cities. Several choices of the number of knots were considered, one knot per 1, 2, 3, 6 and 12 months. This approach resulted in a selection of one knot for every 2 months. Natural splines terms for temperature and pressure using 2 knots or 4 degrees of freedom were included in the model. A 0.50 span resulted in approximately 2 degrees of freedom for both the weather variables. All natural spline models contained day-of-the-week indicator functions.

The degree of complexity of the time model needed to drive the model residuals to white noise was much greater when all days of observation were included in the analysis. However, far fewer degrees of freedom were required to produce white noise residuals when using only those days in which particulate measurements were examined. The every sixth day mortality data displayed more variable temporal structure than the every-day data. The lack of consistency in temporal structure led to a model with fewer degrees of freedom. That is, the degree of freedom needed to adequately fit the temporal trends in daily mortality data clearly overfit the six-day mortality data. These models were rerun with days for which particulate data were available.

Alternate methods of model selection were considered. The AIC was determined for each model in addition to the sum of the AICs among the 8 cities for time degrees of freedom equivalent to a knot per 1, 2, 3, 6, and 12 months. The knot per month value that resulted in the smallest sum of AICs across cities was selected. Two modeling criteria for selecting a common number of knots among all cities were thus considered.

Finally, a knot selection method was examined that allowed the number of knots to vary by city. There, the number of knots was determined for each city separately based on the white noise or AIC criteria. Thus, eight different sets of analyses based on three factors were examined: Common versus City-Specific; AIC versus White noise; and All Days versus Particulate (PM) Days.

REVISED ANALYSIS RESULTS

The pooled estimates ($\times$ 1000) of $PM_{2.5}$, $PM_{10-2.5}$, and PM_{10} associations with nonaccidental mortality among the eight cities and the corresponding t values (ratio of effect estimate to standard error) are given in Table 1 for models using LOESS by inclusion of day-of-the-week indicator variables and convergence criteria. These values are approximately the percentage of change in daily mortality associated with a 10 µg/m^3 change in mass. The size of the pooled effect estimates increased after including day of the week while the estimates decreased when the stricter convergence criteria were used.

The magnitude of the association between $PM_{2.5}$ and mortality varied by the number of knots of the natural spline for time: 1.17% (t = 3.12) for knot/year, 1.10% (t = 2.28) for knot/6 months, 0.86% (t = 2.14) for knot/3 months, 0.85% (t = 2.07) for knot/2 months, and 0.75% (t = 1.72) for knot/month. However, a greater variation in effect estimates was observed for $PM_{10-2.5}$: 1.53% (t = 3.42) for knot/year,

Table 1. Pooled Estimate Among Eight Cities of Percentage Change in Daily Mortality Associated with a 10-µg/m^3 Change in Size-Fractionated Particle Mass Measured on the Day Prior to Death. Pooled Estimate Obtained Using LOESS for Time with 90-Day Span, and LOESS for Temperature and Maximum Change in Barometric Pressure, with a 0.5 Span by Convergence Criteria and Inclusion of Day-of-the-Week Indicator Variables. Ratio of Pooled Estimate to Standard Error (t value) Given in Parenthesis.

	Pooled Percentage of Change (t value)			
	Default Criteria[a]		Strict Criteria[b]	
Pollutant	Without DOW	With DOW	Without DOW	With DOW
$PM_{2.5}$	1.35 (4.17)	1.62 (3.83)	1.11 (3.43)	1.44 (3.14)
$PM_{10-2.5}$	0.91 (2.30)	1.00 (2.45)	0.75 (1.89)	0.83 (2.04)
PM_{10}	0.80 (3.69)	0.87 (3.92)	0.64 (2.95)	0.70 (3.15)

[a] Default convergence criteria: relative difference in parametric parameter estimates between local scoring iterations of 10^{-3}; a difference in the residual sum of squares between backfitting iterations of 10^{-3}; maximum iterations for the local scoring algorithm of 10; and maximum iterations for the backfitting algorithm of 10.

[b] Strict convergence criteria: relative difference in parametric parameter estimates between local scoring iterations of 10^{-15}; a difference in the residual sum of squares between backfitting iterations of 10^{-15}; maximum iterations for the local scoring algorithm of 1000; and maximum iterations for the backfitting algorithm of 1000.

1.16% (*t* = 2.52) for knot/6 months, 0.82% (*t* = 1.69) for knot/3 months, 0.73% (*t* = 1.46) for knot/2 months, and 0.49% (*t* = 0.91) for knot/month. These results suggest that the air pollution association with mortality is sensitive to the amount of temporal smoothing. Thus strategies are needed to select the amount of smoothing using some clearly defined criteria.

The pooled effect estimates for the three particulate pollutants are given by the eight knot selection strategies in Table 2. A knot per 2 months was selected based on both White noise and AIC for the All Days data under the Common model. This analysis strategy is identical to the one used in the original publication except natural splines replaced LOESS. The air pollution effect estimates based on natural splines were smaller than those based on LOESS, even for the stricter convergence criteria (0.85% vs 1.44% for PM$_{2.5}$, 0.73% vs 0.83% for PM$_{10-2.5}$, and 0.53% vs 0.70% for PM$_{10}$).

The City-Specific knot selection strategy yielded somewhat smaller effect estimates than the Common strategy while the White noise strategy generally produced smaller effect estimates than the AIC strategy. Larger differences in effect estimates were observed for the PM Days compared with the All Days strategy.

DISCUSSION

The association between size-fractionated particulate mass and nonaccidental mortality in eight of Canada's largest cities over the 11-year period from 1986 to 1996 was sensitive to the method of statistical analysis. Nonparametric LOESS yielded larger risk estimates than parametric natural splines. Nonparametric spline smoothers gave intermediate risk effects (results not shown). The risk estimates generally varied inversely with the number of knots used to define the natural splines of time. PM$_{2.5}$ effects were less sensitive to the number of knots than the effects of PM$_{10-2.5}$ on mortality. This result was due to the fact that the nonlinear correlation, or concurvity, between the modeled temporal trends in mortality and mass was stronger for PM$_{10-2.5}$ (average correlation among cities of -0.45) compared to PM$_{2.5}$ (-0.36). Thus separating the confounding effects of time and PM$_{10-2.5}$ was more difficult than time and PM$_{2.5}$ in this study.

Positive estimates of heterogeneity of particulate effect, $\hat{\theta}$, across cities were observed using LOESS while negative estimates of heterogeneity were obtained using natural splines. This finding was due to the reduction in effect estimate using natural splines that resulted in smaller

Table 2. Pooled Estimate Among Eight Cities of Percentage Change in Daily Mortality Associated with a 10-µg/m^3 Change in Size-Fractionated Particle Mass Measured on the Day Prior to Death. Pooled Estimate Obtained by Selecting Degrees of Freedom (*df*) for Natural Spline of Time (Common/City-Specific, Analysis Days, and Fitting Criteria). Model Also Includes Natural Spline of Daily Average Temperature with 4 *df* and Maximum Change in Barometric Pressure Within a Day with 4 *df*. Ratio of Pooled Estimate to Standard Error (*t* value) Given in Parenthesis.

Common/ City Specific[a]	Analysis Days[b]	Fitting Criteria[c]	Percent Change in Mortality for 10 µg/m^3 Change in Mass		
			PM$_{2.5}$	PM$_{10-2.5}$	PM$_{10}$
Common (knot/year)[d]	PM days	AIC	1.17 (3.12)	1.53 (3.42)	0.95 (3.75)
Common (knot/6 months)	PM days	White noise	1.10 (2.89)	1.16 (2.52)	0.80 (3.10)
City specific	PM days	AIC	0.90 (2.34)	1.42 (3.13)	0.81 (3.12)
City specific	PM days	White noise	1.05 (2.78)	1.21 (2.65)	0.80 (3.11)
Common (knot/2 months)	All days	AIC and white noise	0.85 (2.07)	0.73 (1.46)	0.53 (1.88)
City specific	All days	AIC	0.75 (1.82)	0.78 (1.58)	0.50 (1.77)
City specific	All days	White noise	0.75 (1.81)	0.58 (1.15)	0.43 (1.51)

[a] Common: equal degree of freedom of natural spline model for time for each city; City specific: degree of freedom of natural spline model for time varies by city.

[b] PM Days: degree of freedom of natural model for time based on days in which particulate data was available; All Days: degree of freedom of natural model for time based on all days of observation (4018 per city).

[c] AIC: Akaike information criterion; white noise: model residuals are tested for evidence that they depart from white noise using the Bartlett test.

[d] Frequency of knot placement for natural splines when Common model is used.

observed variation in the effect estimates, $\hat{\beta}_j$, across cities in addition to the increased within-city estimation error $(\hat{\nu}_j)$ compared to models using LOESS for time and weather. Evidence from this study is insufficient to conclude that the particulate association with mortality varies across Canadian cities.

Fewer knots were required to adequately model the temporal variation in mortality when days with particulate data were examined compared to all days of observation of deaths due to the difficulty in modeling small temporal changes within a season using the sparser dataset. Two knots per year were adequate to model a seasonal cycle in mortality with a winter peak and a summer trough. These seasonal cycles could be identified with the PM Days data, but the data were too sparse to recognize subseasonal variations. However, such variations exist in the daily mortality time series as evidenced by the need for more knots to model time in the daily data (knot/2 months). Some positive correspondence between these subseasonal variations in mortality and subseasonal variations in particulate matter was shown by the larger effects observed when a knot/12 months was used versus a knot/2 months.

REFERENCES*

Burnett RT, Brook J, Dan T, Delocla C, Philips O, Cakmak S, Vincent R, Goldberg MS, Krewski D. 2000. Association between particulate- and gas-phase components of urban air pollution and daily mortality in eight Canadian cities. Inhalation Toxicol 12(Suppl 4):15–39.

Dominici F, McDermott, A, Zeger SL, Samet JM. 2002. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiol 156:193–203.

Priestly MB. 1981. Spectral Analysis and Time Series. Academic Press, London, England.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
DOW	day of the week
GAM	generalized additive model
LOESS	locally weighted smoothers
PM	particulate matter
$PM_{2.5}$	PM 2.5 µm in median aerodynamic diameter
PM_{10}	PM 10 µm in median aerodynamic diameter
$PM_{10–2.5}$	PM 10–2.5 µm in median aerodynamic diameter

* Bold type identifies publication containing the original analyses revised in this short communicaiton report.

Shape of the Exposure–Response Relation and Mortality Displacement in the NMMAPS Database

Francesca Dominici, Michael Daniels, Aidan McDermott, Scott L Zeger, and Jonathan M Samet

ABSTRACT

This section summarizes the revised results of three analyses of the National Morbidity, Mortality, and Air Pollution Study (NMMAPS*) data for particulate matter (PM) less than 10 μm in aerodynamic diameter (PM_{10}) and mortality: (1) estimation of shape of the exposure–response relation and location of any threshold among the 20 largest US cities (1987–1994); (2) estimation of regional and national average exposure–response curves among the 88 largest US cities (1987–1994); and (3) frequency and time-domain log-linear regression analyses to assess the extent of mortality displacement in Philadelphia (1974–1987). These analyses did not show any substantial differences from previous analyses.

INTRODUCTION

Recent work by Dominici and colleagues (2002c) and Ramsay and coworkers (2003) called for caution in use of the S-Plus software for fitting generalized additive models (GAMs) (Hastie and Tibshirani 1990) to estimate relative rates of mortality and morbidity in time-series studies of air pollution and health. When data to which GAMs are applied have small estimated regression coefficients (relative to the effects of confounders) and when confounding factors are modeled with nonparametric smooth functions, the defaults in the S-Plus software (version 3.4) for GAM do not assure convergence of its iterative estimation procedure (Dominici et al 2002c; T Hastie, personal communication, 2003). Thus, biased estimates of the regression coefficients and their standard errors can result. Independently, Ramsay and coworkers (2003) noted that the S-Plus GAM

function uses a computational approximation that, when the air pollution variable correlates with the nonlinear functions included in the model, can underestimate standard errors of the relative rates. (See also Chambers and Hastie [1992] pages 303–304 and commentaries by Lumley and Sheppard [2003] and Samet et al [2003].) Standard errors are underestimated even when more stringent convergence parameters are used.

Recently, some progress has been made to overcome limitations of the S-Plus GAM software. First, default convergence parameters of the S-Plus GAM function (version 6.1) have been made more stringent with a change from 10^{-3}, the default, to 10^{-7}. Second, a revised version of the S-Plus GAM calculates asymptotically exact standard errors of the relative rate estimate (see Dominici et al [2002b] for reference and *http:// biosun01.biostat.jhsph.edu/~fdominic /research.html* for software). In this short communication, we present findings of new analyses of three NMMAPS databases: (1) estimation of a national average dose-response curve for the largest 20 US cities (Daniels et al 2000); (2) estimation of a national average dose-response curve for the largest 88 cities (Dominici et al 2002a); and (3) frequency- and time-domain log-linear regression to assess the extent of mortality displacement (Zeger et al 1999). All other reanalyses of the NMMAPS Study are summarized in a separate report to the Health Effects Institute submitted in October 2002 (Dominici et al, this volume).

METHODS

ESTIMATION OF DOSE RESPONSE

In the original NMMAPS analyses (Daniels et al 2000; Dominici et al 2002a), we extended the city-specific regression model (Dominici et al 2000) by assuming that the expected value of mortality is a natural cubic spline of air pollution adjusted by several confounders. The confounders included smooth functions of time and temperature, originally modeled as smoothing splines (Kelsall et al 1997; Dominici et al 2000). Within each city, the shape of

---

the exposure-response curve was estimated by using the GAM software with default convergence parameters.

Here we re-estimate the shape of the exposure-response curve using the same approach as in Daniels and associates (2000) and Dominici and colleagues (2002a), but modeling the smooth functions of time and temperature as natural cubic splines and using generalized linear models (GLM; McCullagh and Nelder 1989) in S-Plus.

More specifically, to estimate the shape of the exposure response, we modeled the logarithm of expected value of daily mortality (E) as a smooth function of air pollution after adjusting for other confounders:

$$E(Y_t) = \exp\left\{S(X_t, \text{knots} = \nu) + \text{confounders}\right\} \qquad (1)$$

where X_t and Y_t are the air pollution and mortality time series, and $S(.,.)$ is a natural cubic spline with fixed boundary knots, with knots (k) at locations $\nu = (\nu_1, \ldots, \nu_k)$. In addition, to examine whether the health effects of air pollution are negligible below some level, a linear threshold model can be used:

$$E(Y_t) = \exp\left\{\theta(X_{t-1} - h)^+ + \text{confounders}\right\} \qquad (2)$$

where ($x^+ = x$ if $x \geq 0$ and $x^+ = 0$ if $x < 0$) and h is an unknown change-point that is estimated from the data. In both modeling approaches, the term *confounders* represents time-varying covariates (as, for example, long-term trends and seasonality in the mortality time-series and weather variables), which may bias the particle–mortality association. These confounders are modeled as natural cubic splines (Table 1).

MORTALITY DISPLACEMENT

The original work by Zeger and colleagues (1999) developed and illustrated an approach for estimating the association between air pollution and mortality from time-series data that is resistant to short-term harvesting. The method was based on the concept that harvesting alone creates associations only at shorter-time scales. It used frequency-domain log-linear regression (FDLLR) (Kelsall et al 1999) to decompose information about the pollution–mortality association into distinct time scales and then created harvesting-resistant estimates by excluding the short-term information that is affected by harvesting. This method was applied to total suspended particles (TSP) and mortality counts from Philadelphia for 1974 through 1988 and showed that the TSP–mortality association in Philadelphia was inconsistent with the harvesting-only hypothesis and that the harvesting-resistant estimates of the TSP relative risk were actually larger, not smaller, than the ordinary estimates. Because the software for fitting FDLLR used GAM with default convergence parameters, reanalyses are now necessary.

Here we reanalyzed the Philadelphia data using two methods: (1) applying the FDLLR software but with stringent convergence parameters (10^{-15}) for the S-Plus function GAM; and (2) developing an alternative method based on the time-domain analog of FDLLR, which uses natural cubic splines and GLM. Time-domain log-linear regression (TDLLR) is based on a Fourier decomposition of air pollution time series into a set of independent exposure variables, each representing a different time scale. These variables are then used as predictors in a Poisson regression model to estimate a separate relative rate of mortality on each exposure time scale while controlling for time and temperature. More specifically, let X_t^c be the air pollution time series, and Y_t^c be the mortality time series in location c. We first decompose the air pollution

Table 1. Potential Confounding Factors or Predictors in Estimation of City-Specific Relative Rates Associated with Particulate Air Pollution Levels and Rationale for Their Inclusion in the Model

Predictors	Primary Reasons for Inclusion
Indicator variables for the three age groups	Allow different baseline mortality rate within each age group
Indicator variables for the day of the week	Allow different baseline mortality rate within each day of the week
Natural cubic splines of time with 7 *df* per year	To adjust for long term trend and seasonality
Natural cubic splines of temperature with 6 *df*	To control for the known effects of weather on mortality
Natural cubic splines of dewpoint with 3 *df*	To control for the known effects of humidity on mortality
Separate natural cubic splines of time (2 *df* for year) for each age group	To separately adjust for seasonality within each age goup

series X_t^c into distinct component series X_{kt}^c, one for each distinct time scale k, and then we calculate the association between Y_t^c without decomposition, and each of the time scale components X_{kt}^c. The decomposition is obtained by applying the discrete Fourier transform to the X_t^c series (Bloomfield 1976). Specifically, we assume:

$$E(Y_t^c) = \exp\left\{ \sum_k \beta_k^c X_{kt}^c + \text{confounders} \right\} \quad (3)$$

where β_k^c (the parameter of interest) denotes the log-relative rate of daily mortality for each 10 unit increase in the air pollution level in location c on a time scale k. The TDLLR and its application to four NMMAPS cities are described in detail elsewhere (Dominici et al 2003).

RESULTS

EXPOSURE RESPONSE RELATION

Differences in the shape of the exposure–response curves were not observed when comparing the new with the original analyses. Figure 1 shows exposure–response curves for total (TOTAL) mortality from nonexternal causes, mortality from cardiovascular and respiratory (CVDRESP) causes, and other causes (OTHER) for the 20 largest US cities, 1987–1994. Figure 2 shows regional PM_{10}-mortality exposure–response curves for TOTAL, for each region, and the national average, for the 88 largest US cities, 1987–1994. The national average exposure–response curves for both

the 20 and 88 cities are linear, and the posterior probability of zero knots is approximately 1. At the regional level, the data from cities in several regions (Northwest, Upper Midwest, and Southeast) indicate some modest departure from a linear model. In particular, the Northwest and Upper Midwest regions show a leveling (saturation effect) at higher levels of PM. However, the uncertainty boundaries for these regions indicate compatibility of the data with a linear relation, and we cannot explain why these regions might have other than a linear exposure–response curve.

Figure 3 shows posterior probabilities for a threshold for the effect of PM_{10} on the cause-specific mortality groupings for the 20 largest US cities, 1987–1994. The posterior distributions on the location of the threshold are skewed to the right for TOTAL and CVDRESP and are skewed to the left for OTHER. The category of OTHER had the highest, most-probable threshold, at 65 µg/m^3. For CVDRESP, however, the data give more support to a low value of h (as 0, 5, 10 µg/m^3), lower than TOTAL. The threshold, if any, for CVDRESP may be lower than for TOTAL.

MORTALITY DISPLACEMENT

Figure 4 shows the frequency-domain estimate of the relative rate for mortality associated with air pollution as a function of frequency. (The frequency-domain estimates were derived from a GAM model that included current-day pollution level, a smooth function of time with a total of 90 df, a smooth function of current-day temperature with 6 df, and a smooth function of current-day dew point temperature with 6 df.) The horizontal axes denote the Fourier frequencies (bottom) and the time scale in days
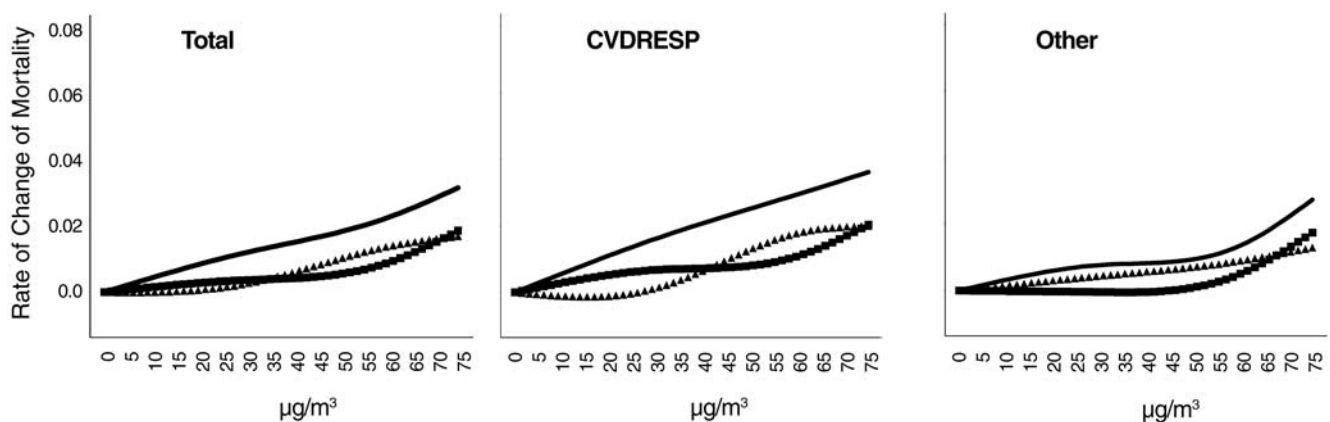
Figure 1. Mortality–PM$_{10}$ dose-response curves for total nonaccidental mortality (Total), cardiovascular and respiratory mortality (CVDRESP), and mortality from other causes (Other), 20 largest US cities, 1987–1994. The exposure-response curves for the mean lag, current day, and previous day PM$_{10}$ are denoted by solid lines, squared points, and triangle points, respectively.
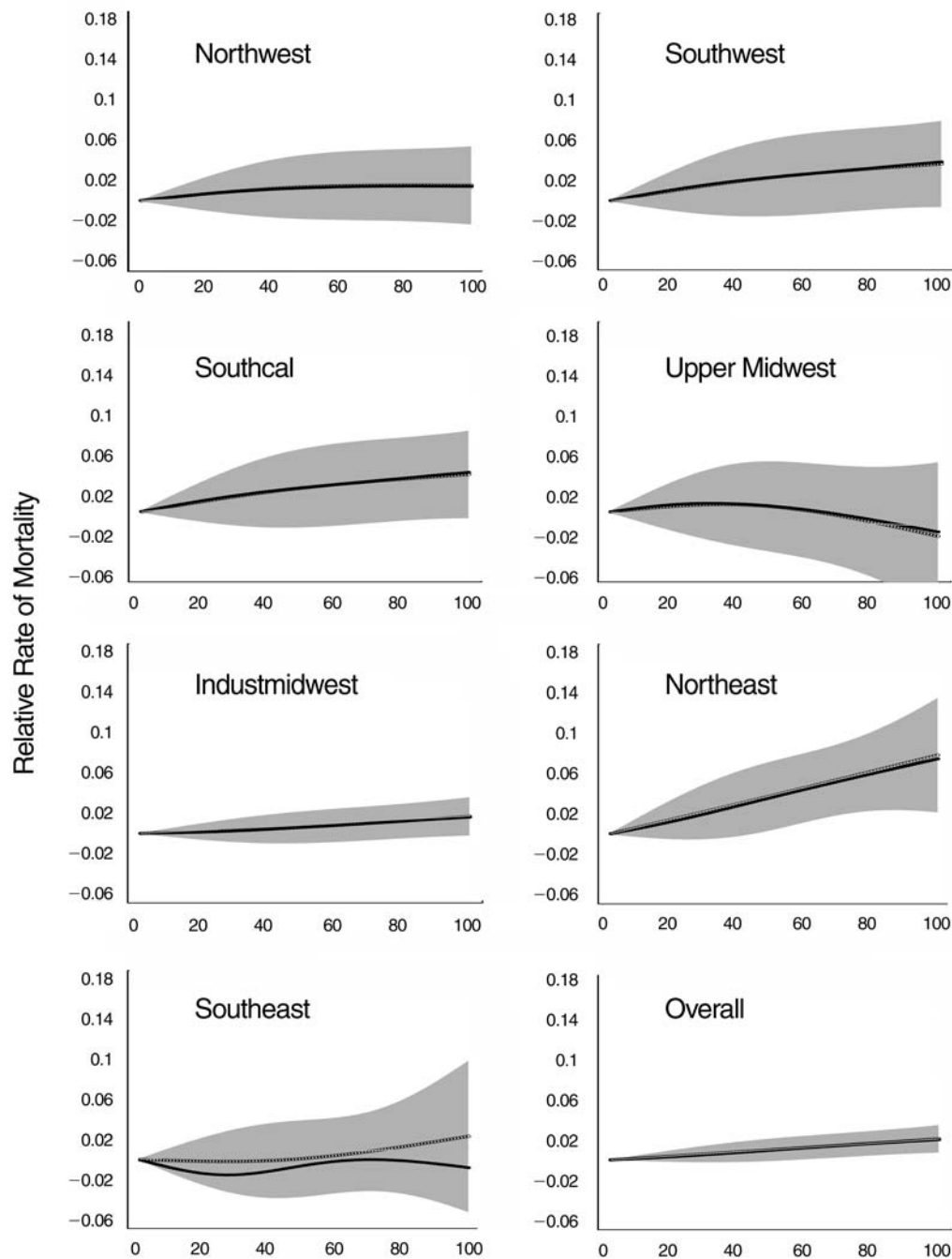
Figure 2. Regional and national-average PM$_{10}$-mortality exposure-response curves. Solid black curves are obtained by fitting the spline model with the reversible jump Markov chain Monte Carlo (RJMCMC) (Green 1995) and allowing for an unknown number and location of knots. The curves with the empty dots are obtained by setting one knot at 40 μg/m^3 and fitting the spline model with a Gibbs sampler. The linear curves are obtained by fitting the hierarchical linear model with a Gibbs sampler without borrowing strength across regions. The shaded area denotes the 95% confidence bands for the curve with a fixed knot.
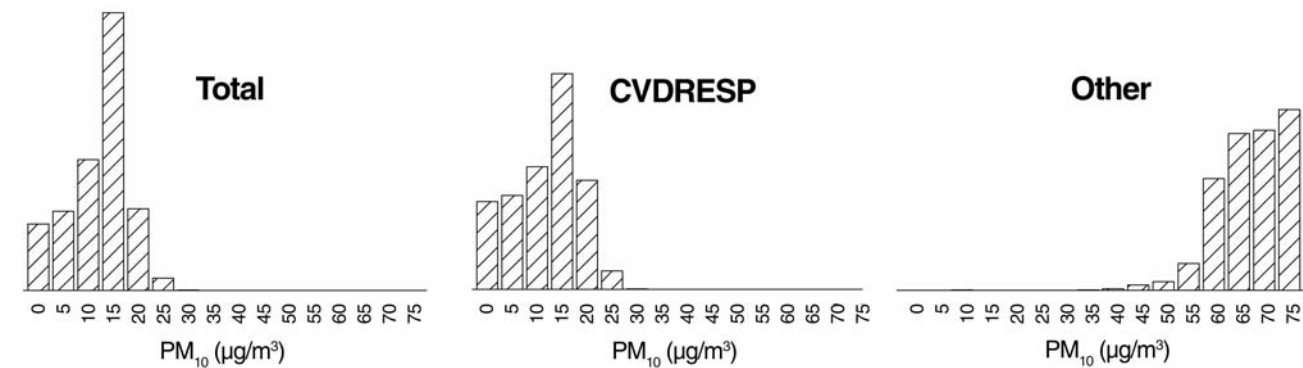
Figure 3. Posterior probabilities of thresholds for each cause-specific mortality and for mean lag PM$_{10}$, 20 largest US cities, 1987–1994. Total, total nonaccidental mortality; CVDRESP, cardiovascular mortality and respiratory mortality; Other, mortality from other causes.
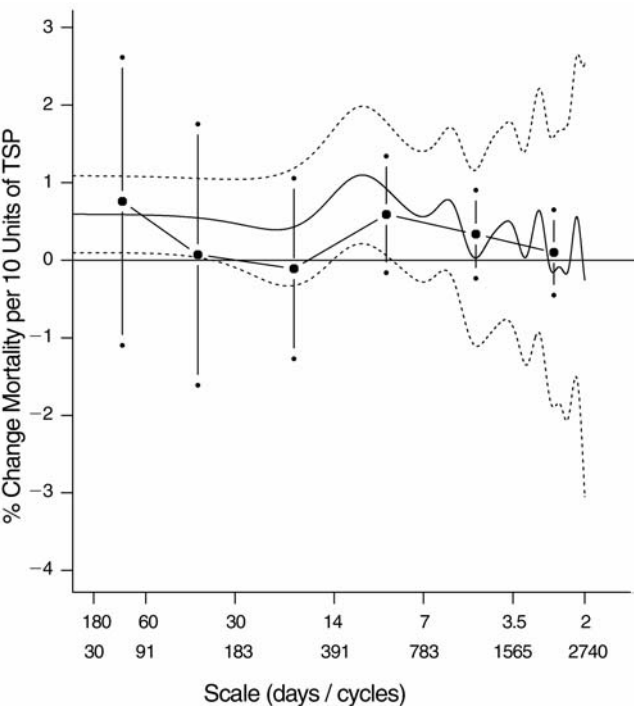
Figure 4. Philadelphia database 1974–1988: comparison between frequency domain (continuous curve) and time-scale estimates (points connected by line segments), showing the log-relative rates of Total mortality by frequency and frequency grouping. The dotted lines show ± 2 SE for the frequency domain estimates and the bars represent ± 2 SE for the time-scale estimates. TSP = total suspended particles.

(top) at which the association is measured. The dotted curves denote the estimated relative rates, plus or minus two estimated standard errors, respectively, at each fre-

quency. The time-scale estimates (points connected by line segments) are plotted on top of the frequency-domain results (continuous curve). Time-scale and frequency-domain results are similar, indicating little sensitivity of the results to the modeling assumptions. In addition, consistent with results for the four NMMAPS cities (Dominici et al 2003), relative rate estimates at longer time-scales are larger than relative rate estimates at shorter time scales, indicating a pattern of air pollution effects at different time scales that is inconsistent with the harvesting-only hypothesis.

DISCUSSION

We have presented NMMAPS reanalyses of the shape of the exposure–response curves for the 20 and 88 cities applying FDLLR and TDLLR to assess mortality displacement.

The exposure–response analyses were performed by using the same modeling approaches described in previous publications but replacing the smoothing spline with natural cubic splines with the same degrees of freedom for the adjustment of time-varying confounders. On average, the shape of the exposure–response curve is linear, confirming results reported in the original analyses.

The analyses of mortality displacement were performed by using FDLLR software with GAM with stringent convergence parameters and by developing and applying an alternative time domain approach (TDLLR) (Dominici et al 2003), which used natural cubic splines and GLM software. As reported in previous analyses, both methods produce patterns of relative rate estimates at different time

scales that are inconsistent with the hypothesis of mortality displacement.

REFERENCES*

Bloomfield P. 1976. Fourier Analysis of Time Series: An Introduction. John Wiley & Sons, New York NY.

Chambers J, Hastie T. 1992. Statistical Models in S. Chapman and Hall, London, England.

Daniels MJ, Dominici F, Samet JM, Zeger SL. 2000. Estimating particulate matter-mortality dose-response curves and threshold levels: An analysis of daily time-series for the 20 largest US cities. Am J Epidemiol 152:397–406.

Dominici F, Daniels M, Zeger SL, Samet JM. 2002a. Air pollution and mortality: Estimating regional and national dose-response relationships. J Am Stat Assoc 97:100–111.

Dominici F, McDermott A, Hastie T. 2002b. Semiparametric Regression in Time Series Analyses of Air Pollution and Health. Technical Report. Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore MD.

Dominici F, McDermott A, Zeger SL, Samet JM. 2002c. On generalized additive models in time series studies of air pollution and health. Am J Epidemiol 156:1–11.

Dominici F, McDermott A, Zeger SL, Samet JM. 2003. Airborne particulate matter and mortality: Time-scale effects in four US cities. Am J Epidemiol. In press.

Dominici F, Samet JM, Zeger SL. 2000. Combining evidence on air pollution and daily mortality from the twenty largest US cities: A hierarchical modeling strategy (with discussion). J R Stat Soc Ser A 163:263–302.

Green PJ. 1995. Reversible jump markov chain monte carlo computation and bayesian model determination. Biometrika 82:711–732.

Hastie TJ, Tibshirani RJ. 1990. Generalized Additive Models. Chapman and Hall, London, England.

Kelsall J, Samet JM, Zeger SL. 1997. Air pollution, and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Kelsall J, Zeger SL, Samet JM. 1999. Frequency domain log-linear models; air pollution and mortality. J R Stat Soc Ser C 48:331–344.

Lumley T, Sheppard L. 2003. Time series analyses of air pollution and health: Straining at gnats and swallowing camels? Epidemiology 14:13–14.

McCullagh P, Nelder JA. 1989. Generalized Linear Models (second edition). Chapman and Hall, London, England.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Samet J, Dominici F, McDermott A, Zeger SL. 2003. New problems for an old design: Time series analyses of air pollution and health. Epidemiology 14:11–12.

Zeger SL, Dominici F, Samet J. 1999. Harvesting-resistant estimates of air pollution effects on mortality. Epidemiology 10:171–175.

ABBREVIATIONS AND OTHER TERMS

CDVRSP	cardiovascular and respiratory mortality
FDLLR	frequency-domain log-linear regression
GAM	generalized additive model
GLM	generalized linear model
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
PM_{10}	particulate matter of 10 μm mass median aerodynamic diameter
TDLLR	time-domain log-linear regression approach
TSP	total suspended particles

* Bold type identifies publications containing the original analyses revised in this short communication report.

Mortality and Air Pollution for Santa Clara County, California, 1989–1996

David Fairley

ABSTRACT

The data are reanalyzed using generalized additive models (GAMs*) with stricter convergence criteria. The results are essentially unchanged with coefficients differing by at most ± 0.5 standard error (SE) from the original and relative risks differing by at most ± 0.01 SE. Various inference methods are compared. The method comparing GAM coefficients with standard errors generated from generalized linear models (GLMs) produces P values similar to simulations. The S-Plus ANOVA (analysis of variance) feature often gave conservative results. As found previously, a statistically significant relation existed between daily nonaccidental mortality and every criteria pollutant, either on the same day or lag one. Particulate matter (PM) less than 2.5 μm in diameter ($PM_{2.5}$) and nitrate (NO_3) predominate when included in models with carbon monoxide (CO), nitrogen dioxide (NO_2), and sulfate (SO_4). Coarse-fraction PM, less than 10 μm in diameter (PM_{10}), was not statistically significant. A new ozone (O_3) variable—the daily number of parts per billion (ppb)-hours greater than a 60-ppb threshold (o3ppbgt60)—was found to have a statistically significant relation with nonaccidental mortality even when included in a regression jointly with $PM_{2.5}$ or NO_3 and was also significantly related to cardiovascular mortality.

INTRODUCTION

> Ockham's Razor: A rule in science and philosophy that… the simplest of two or more competing theories is preferable and that an explanation for unknown phenomena should be first attempted in terms of what is already known.
>
> —*The American Heritage Dictionary*, 2nd Edition

Problems have recently been reported in use of the GAM function in the S-Plus statistics package (Dominici et al 2002). The GAM procedure uses an iterative algorithm to find the solution. Further analysis found that the default number of iterations and the stopping criterion sometimes produced results that were far from those obtained after the algorithm converged. Furthermore, the standard errors provided in S-Plus were shown to underestimate the true standard errors. Questions about the results using GAM are of particular concern because the current reevaluation of the national particulate standards are based on time-series studies, many of which used this function (for example, Dominici et al 2000; Kelsall et al 1997; Moolgavkar 2000; Saez et al 2002; Samet et al 2000; Schwartz 1994).

Because of questions raised by these problems, we decided to reanalyze the data from our previous study (Fairley 1999). In our previous analysis, we used GAM but with a more stringent stopping criterion than the default so that the results in this reanalysis changed only in a minor way. Moreover, the previous analysis did not use the standard errors provided in the S-Plus summary GLM function. Instead we used an F test from the ANOVA function, which is based on the change in −2 log (likelihood), to determine whether the addition of a variable was statistically significant. Thus, conclusions about statistical significance are also essentially unchanged.

This analysis contained some innovations. One was to augment the $PM_{2.5}$ data using other variables. A second was to use an alterative ozone variable—o3ppbgt60.

DATA AND METHODS

The same data and variables were used as in the previous study with two additions (o3ppbgt60 and pm2.5aug [$PM_{2.5}$ measurements augmented with $PM_{2.5}$ predicted from COH and PM_{10}]): 24-hour $PM_{2.5}$, PM_{10}, $PM_{10-2.5}$, NO_3 and SO_4; 24-hour averaged coefficient of haze (COH) and NO_2; and 8-hour averaged CO and O_3, for 1989 through 1996 from the

4th Street monitoring site in San Jose Santa Clara County, California. Regressions were performed for all nonaccidental deaths to residents who died in the county: *International Classification of Diseases* (ICD) codes 0–799; respiratory mortality (ICD codes 11, 35, 472–519, 710.0, 710.2, 710.4); and cardiovascular mortality (ICD codes 390–459).

The same methods were used as in the previous study except for increasing the stringency of the stopping rule. Briefly, a Poisson regression model was fit with GAM terms for day of year and trend with the best fit determined by minimizing Akaike information criterion (AIC). GAM terms for minimum and maximum temperatures were added to this model and again the model with the minimum AIC was determined. The degrees of the smooths were the same as found previously. Various pollutant terms were then added to this best fitting model.

The stopping rule was tightened with epsilon and back-fit epsilon set to 10^{-12}, whereas previously it had been 10^{-4} (already been more stringent than the S-Plus default, 10^{-3}) in the previous analysis. The maximum number of iterations was raised from the default, 10, to 10^7. These were chosen based on trial and error so that there was no change if the stopping rule was made more stringent.

For comparison, a parallel analysis was done using GLM regression. Here, natural splines were used in place of smoothing splines. AIC was again used to find the optimal degree of freedom (*df*) for day of year, trend, and minimum and maximum temperatures. The same strict convergence criteria were used as with the GAM regression.

The reanalysis updates Tables 4, 5 and 6 from Fairley (1999). As mentioned previously, inference on whether a variable included in the fit was statistically significant was based on the deviance [change in $-2 \log(\text{likelihood})$] from adding that variable, using the ANOVA feature in S-Plus.

S-Plus does not provide standard errors for GAM regressions and advises using a suggestion from Chambers and Hastie (1992): "In practice, one can always approximate the nonparametric term parametrically (and even conservatively) using functions such as bs () [B spline] or ns () [natural spline], and use the inexpensive parametric standard-error curves." Following this advice, standard error estimates were obtained by running GLM regressions, replacing smoothing splines with natural splines of the same degree. These standard errors were used to compare the differences in coefficients between this and the original analysis and between GAM and GLM regressions.

To check the reliability of these inferences, simulations were performed from the model fit without pollutant variables (that is, assuming that the null-model fitted parameters were the true parameters), generating Poisson variates from this model and then refitting the model including a pollutant variable. This simulated the null distribution of the pollutant coefficient, allowing comparison of the pollutant coefficient estimated using the actual data. A more general simulation was also rerun for $PM_{2.5}$ and pm2.5aug. Here the whole model-building process is simulated—finding the model with the best AIC for time/season terms, then finding the model with the best AIC for temperature, and finally adding the PM term. See Fairley (1999) for more details. The only difference with the 1999 simulations was using epsilons of 10^{-12} and 10^3 iterations.

As a third comparison, inferences were based on the ratio of the fitted GAM coefficients to the GLM standard errors mentioned previously.

Note that some errors were found in the original analysis. One was the miscoding of missing ozone values. The other two were mistaken entries. One mistake was the relative risk of COH with respiratory mortality. The relative risk (RR) 1.07 should have been 1.10 (and highly significant). The second mistake was the RR for $PM_{2.5}$ for spring seasonal regression, which was listed in the original document as 1.05 when it should have been 0.92.

The original dataset had only 408 $PM_{2.5}$ observations so the power was borderline for detecting an effect. We found that $PM_{2.5}$ could be well predicted from PM_{10} and COH, more than doubling of the number of $PM_{2.5}$ values. In particular, a linear regression yielded a fit of $PM_{2.5} = 0.392 \times PM_{10} + 14.0 \times COH - 4.02$, with a regression standard error of 4.49 µg/m^3 and a multivariate coefficient of determination (R^2) of 0.88. The total number of pm2.5aug observations was 835.

A second and perhaps more significant innovation was to use an alternate variable for ozone. Previously, 8-hour maximum ozone (oz8hr) was used because it corresponded to the national standard. However, the 8-hour averages often contain very low ozone values; clinical studies have not found ozone health effects below about 80 ppb and the natural background is approximately 40 ppb. Thus, it seemed reasonable to consider the daily number of ppb-hours above a threshold, with thresholds between 40 ppb and 80 ppb. The highest correlations with daily mortality were found using a threshold of 60 ppb, so those results are reported here. Figure 1 shows this variable plotted against daily mortality. This variable, denoted *o3ppbgt60*, is highly skewed. The largest values are clearly influential observations.
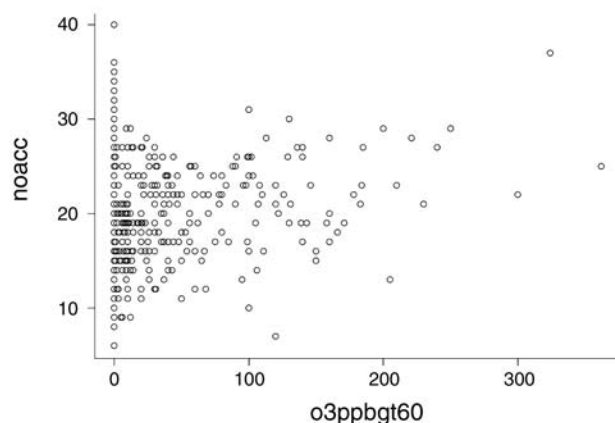
Figure 1. Daily nonaccidental deaths versus daily ozone ppb-hours greater than 60 ppb.

RESULTS

COMPARISON OF RESULTS FROM STRICT AND DEFAULT GAM ANALYSIS

Tables 1a and 1b present RRs from the old and new GAM regressions for all nonaccidental cause mortality. To compare the magnitude of the changes in coefficients, the GLM standard errors were used to normalize the differences (that is, taking the ratio of the difference between the coefficient from the old and new analyses to the standard error). There were 34 analyses that were repeated in Tables 1a and 1b. The range of normalized differences was −0.34 to +0.40, but 40 were between −0.09 and +0.05.

The relative risk estimates remained virtually identical, changing by at most ± 0.01. The statistical significance changed in only one case: same-day oz8hr was not found to be statistically significant in the new analysis (although the estimated coefficient increased somewhat). The difference was not due to GAM but to the miscoding of ozone data in the original analysis (as mentioned earlier).

The results for cardiovascular and respiratory mortality in Table 2 are almost identical to those in the original analysis. The one exception, as mentioned previously, is that the RR for lag COH was 1.10, which is highly significant; the previously reported RR of 1.07 was an error. The pm2.5aug data were statistically significant in the regression with cardiovascular mortality, whereas the coefficient for the nonaugmented $PM_{2.5}$ was not.

The RRs for Table 3 are identical except, as mentioned before, for $PM_{2.5}$ in summer; pm$_{2.5}$aug is statistically significant in the winter months.

COMPARISON OF RESULTS FROM DEFAULT GAM AND GLM ANALYSES

The same 34 values from Table 1 were also computed using a GLM regression. The normalized differences were greater on the whole than between the old and new GAM analyses, with a median absolute difference of 0.15. The differences ranged from −0.18 to +0.37. The GLM coefficients were statistically smaller, with an average normalized difference of 0.081. The GLM results are very similar to the GAM approach, with no RRs differing by more than 0.02 and all but five differing by 0.01 or less. The same coefficients were statistically significant.

Only PM_{10}, $PM_{2.5}$, pm2.5aug, and o3ppbgt60 GLM analyses are shown in Table 2. The results are again essentially unchanged.

Table 3 presents RRs by season. Here the results are unchanged with two exceptions. Under GLM, NO_3 is no longer statistically significant for the summer months, and lag COH becomes statistically significant for the winter months.

RESULTS FOR DAILY NUMBER OF HOURS WITH O_3 GREATER THAN OR EQUAL TO 60 ppb

For nonaccidental mortality, o3ppbgt60 is statistically significant alone and paired with other pollutants. The inclusion of o3ppbgt60 in a model with other pollutants produces almost no change either in the coefficient estimates or the statistical significance of the other pollutants. In contrast, the inclusion of $PM_{2.5}$ or NO_3 with another variable (except for o3ppbgt60 and oz8hr) reduces the magnitude and statistical significance of the coefficient of the other variable. The o3ppbgt60 was statistically significant in the regression with cardiovascular mortality but not respiratory mortality. For the seasonal regressions in Table 3, o3ppbgt60 is statistically significant in summer but not other seasons.

COMPARISON OF *P* VALUES

Because confidence intervals have been used to compare the results from various studies, it is worthwhile to examine the relation between inferences derived directly from the F tests of one model versus another with the indirect inferences derived from *t* tests comparing the coefficients with their estimated standard errors. Table 4 shows a comparison of *P* values for these tests for the inclusion of

Table 1a. Relative Risks[a] (95% Confidence Intervals[b]) for Daily Mortality from Nonaccidental Causes with Pollutant Alone and Lagged[c]

Pollutant	Same-Day Variables					1-Day Lagged Variables				
	Original GAM	New GAM		New GLM		Original GAM	New GAM		New GLM	
PM$_{10}$	0.080	0.078	(0.028,0.131)	0.083	(0.029,0.139)	−0.009	−0.011	(−0.055,0.035)	−0.012	(−0.059,0.038)
PM$_{2.5}$	0.093	0.092	(0.018,0.172)	0.080	(0.016,0.148)	−0.041	−0.042	(−0.101,0.020)	−0.043	(−0.102,0.021)
pm2.5aug	NA	0.083	(0.030,0.138)	0.094	(0.034,0.158)	NA	−0.002	(−0.076,0.077)	−0.004	(−0.129,0.138)
PM$_{10-2.5}$	0.024	0.023	(−0.040,0.091)	0.017	(−0.028,0.064)	−0.022	−0.023	(−0.080,0.038)	−0.029	(−0.100,0.049)
COH	0.026	0.022	(−0.010,0.054)	0.024	(−0.012,0.061)	0.047	0.043	(0.016,0.071)	0.045	(0.017,0.073)
NO$_3$	0.074	0.074	(0.025,0.124)	0.070	(0.024,0.117)	−0.024	−0.025	(−0.070,0.023)	−0.025	(−0.071,0.023)
SO$_4$	0.053	0.053	(0.007,0.101)	0.052	(0.007,0.100)	−0.017	−0.017	(−0.060,0.027)	−0.014	(−0.049,0.022)
NO$_2$	0.035	0.030	(−0.002,0.062)	0.026	(−0.002,0.055)	0.032	0.028	(0.003,0.055)	0.027	(0.002,0.052)
CO	0.017	0.015	(−0.005,0.035)	0.014	(−0.005,0.033)	0.042	0.039	(0.016,0.064)	0.039	(0.016,0.063)
oz8hr	0.030	0.031	(−0.003,0.066)	0.033	(−0.003,0.070)	-0.008	−0.006	(−0.023,0.012)	−0.005	(−0.022,0.011)
o3ppbgt60	NA	0.038	(0.014,0.063)	0.041	(0.015,0.067)	NA	0.005	(−0.019,0.029)	0.002	(−0.007,0.010)

[a] Relative risks calculated by exp(b*Δp) − 1, where b is the pollutant coefficient from the Poisson regression, and Δp = 50 for PM$_{10}$, corresponding to the increment used in the 1997 Criteria Document for particulate matter. For other pollutants, p, the increment was 50 * sd(p)/sd(pm10). For example, sd(pm2.5) = 13, sd(pm10) = 23, so for pm2.5, Δp = 50 * 13/23 = 28.

[b] The confidence intervals are c ± 2*s, where for GAM, s = abs(c)/sqrt(F), the F obtained from applying ANOVA to the models with and without the pollutant variable, and for GLM, s = standard error provided by S-Plus.

[c] GAMs include a 7 *df* smoothing spline for trend, 12 *df* smoothing spline for season, 3 *df* smoothing spline for minimum temperature and 2 *df* smoothing spline for maximum temperature. GLMs include 7 *df* natural spline for trend, 4 *df* natural spline for day of year, 3 *df* natural spline for minimum temperature and 3 *df* natural spline for maximum temperature. All models include 1 pollutant term entered linearly.

Table 1b. Relative Risks[a] (95% Confidence Intervals[b]) for Daily Mortality from Nonaccidental Causes for $PM_{2.5}$ and pm2.5aug with Other Pollutants

Pollutant	Original GAM	New GAM	New GLM	New GAM	New GLM
		$PM_{2.5}$ with Other Pollutant		**pm2.5aug with Other Pollutant**	
PM_{10}	0.126	0.125 (−0.055,0.340)	0.129 (−0.057,0.352)	0.101 (−0.054,0.283)	0.123 (−0.065,0.348)
$PM_{10-2.5}$	0.111	0.109 (0.021,0.205)	0.099 (0.019,0.185)	0.135 (0.026,0.255)	0.122 (0.023,0.229)
lag COH	0.108	0.107 (0.008,0.217)	0.090 (0.006,0.182)	0.101 (0.027,0.181)	0.105 (0.027,0.188)
NO_3	−0.001	−0.003 (−0.046,0.042)	−0.010 (−0.153,0.157)	0.045 (−0.069,0.173)	0.062 (−0.093,0.244)
SO_4	0.099	0.097 (0.008,0.194)	0.092 (0.007,0.184)	0.082 (−0.053,0.236)	0.087 (−0.055,0.251)
NO_2	0.119	0.119 (0.030,0.216)	0.108 (0.027,0.196)	0.103 (0.037,0.173)	0.111 (0.039,0.187)
lag CO	0.107	0.107 (0.024,0.195)	0.092 (0.021,0.167)	0.091 (0.028,0.159)	0.095 (0.029,0.166)
oz8hr	0.100	0.100 (0.024,0.181)	0.088 (0.021,0.159)	0.087 (0.032,0.145)	0.088 (0.033,0.147)
o3ppbgt60	NA	0.092 (0.008,0.182)	0.070 (0.006,0.139)	0.079 (0.024,0.137)	0.080 (0.024,0.139)
		Other Pollutant with $PM_{2.5}$		**Other Pollutant with pm2.5aug**	
PM_{10}	−0.038	−0.037 (−0.276,0.280)	−0.056 (−0.385,0.450)	−0.015 (−0.136,0.122)	−0.029 (−0.242,0.244)
$PM_{10-2.5}$	−0.030	−0.029 (−0.110,0.059)	−0.031 (−0.117,0.063)	−0.029 (−0.110,0.059)	−0.031 (−0.117,0.063)
lag COH	−0.009	−0.010 (−0.103,0.091)	−0.006 (−0.062,0.053)	−0.017 (−0.074,0.044)	−0.017 (−0.076,0.045)
NO_3	0.089	0.090 (−0.009,0.200)	0.086 (−0.009,0.190)	0.052 (−0.015,0.124)	0.041 (−0.012,0.096)
SO_4	0.001	0.002 (−0.059,0.066)	−0.007 (−0.234,0.287)	0.029 (−0.020,0.081)	0.028 (−0.019,0.076)
NO_2	−0.040	−0.041 (−0.114,0.037)	−0.042 (−0.116,0.038)	−0.022 (−0.067,0.026)	−0.027 (−0.083,0.032)
lag CO	−0.024	−0.025 (−0.082,0.036)	−0.020 (−0.066,0.029)	−0.011 (−0.054,0.034)	−0.012 (−0.058,0.036)
oz8hr	0.038	0.057 (−0.046,0.171)	0.054 (−0.044,0.162)	0.030 (−0.054,0.121)	0.018 (−0.033,0.072)
o3ppbgt60	NA	0.054 (0.005,0.107)	0.073 (0.006,0.144)	0.050 (0.005,0.097)	0.052 (0.005,0.101)

[a] Relative risks calculated by exp(b*Δp) −1, where b is the pollutant coefficient from the Poisson regression, and Δp = 50 for PM_{10}, corresponding to the increment used in the 1997 Criteria Document for particulate matter. For other pollutants, p, the increment was 50 * sd(p)/sd(pm10). For example, sd(pm2.5) = 13, sd(pm10) = 23, so for pm2.5, Δp = 50 * 13/23 = 28.

[b] The confidence intervals are c ± 2*s, where for GAM, s = abs(c)/sqrt(F), the F obtained from applying ANOVA to the models with and without the pollutant variable, and for GLM, s = standard error provided by S-Plus.

Table 2. Relative Risks[a] (95% Confidence Intervals[b]) for Daily Mortality from Cardiovascular and Respiratory Causes with Pollutant Alone[c]

Pollutant	Cardiovascular Mortality[d]			Respiratory Mortality[e]		
	Original GAM[f]	New GAM[f]	New GLM	Original GAM[f]	New GAM[f]	New GLM
PM$_{10}$	0.086	0.085 (0.006,0.170)	0.089 (0.013,0.170)	0.108	0.107 (−0.037,0.272)	0.108 (−0.034,0.271)
PM$_{2.5}$	0.073	0.072 (−0.046,0.205)	0.076 (−0.028,0.191)	0.133	0.133 (−0.110,0.442)	0.154 (−0.041,0.389)
pm2.5aug		0.103 (0.018,0.196)	0.104 (0.025,0.190)		0.102 (−0.033,0.255)	0.108 (−0.038,0.277)
PM$_{10-2.5}$	0.026	0.026 (−0.072,0.134)		0.156	0.156 (−0.065,0.428)	
COH	0.030	0.027 (−0.015,0.072)		0.068[g]	0.097 (0.020,0.180)	
NO$_3$	0.093	0.092 (0.017,0.173)		0.100	0.100 (−0.048,0.272)	
SO$_4$	0.040	0.040 (−0.028,0.113)		0.147	0.147 (0.011,0.301)	
NO$_2$	0.023	0.019 (−0.024,0.065)		0.072	0.069 (−0.004,0.147)	
CO	0.041	0.039 (0.001,0.077)		0.082	0.079 (0.012,0.151)	
oz8hr	0.024	0.026 (−0.023,0.078)		−0.043	−0.047 (−0.122,0.034)	
o3ppbgt60		0.043 (0.004,0.083)	0.047 (0.009,0.086)		−0.018 (−0.071,0.038)	−0.027 (−0.097,0.049)

[a] Relative risks calculated by exp(b*Δp) −1, where b is the pollutant coefficient from the Poisson regression, and Δp = 50 for PM$_{10}$ and 50 * sd(p)/sd(pm10) for other pollutants, p. For example, sd(pm2.5) = 13, sd(pm10) = 23, so for pm2.5, Δp = 50 * 13/23 = 28. Single asterisks indicate statistical significance at the 0.05 level, double asterisks at the 0.01 level.

[b] The confidence intervals are c ± 2*s, where for GAM, s = abs(c)/sqrt(F), the F obtained from applying ANOVA to the models with and without the pollutant variable, and for GLM, s = standard error provided by S-Plus.

[c] GAMs include a 7 df smoothing spline for trend, 12 df smoothing spline for season, 3 df smoothing spline for minimum temperature and 2 df smoothing spline for maximum temperature. GLMs have 7 df natural spline for trend, 4 df natural spline for season, 3 df natural spline for minimum temperature and 3 df natural spline for maximum temperature.

[d] ICD categories 390−459.

[e] ICD categories 11, 35, 472-519, 710.0, 710.2, 710.4.

[f] Original GAM used epsilon and bf.epsilon of 10^{-4} with default # of iterations. New GAM used epsilons of 10^{-12}, and iterations and bf.iterations of 10^7.

[g] Error in the original article; this value should have been 0.100.

Table 3. Relative Risks (95% Confidence Intervals[a]) for Daily Mortality from Nonaccidental Causes by Season[b,c]

	Old GAM	New GAM	New GLM
Spring[d]			
PM_{10}	0.076	0.076 $(-0.086, 0.267)$	0.100 $(-0.045, 0.267)$
$PM_{2.5}$	0.071	0.071 $(-0.132, 0.322)$	0.047 $(-0.132, 0.263)$
pm2.5aug		0.065 $(-0.124, 0.293)$	0.108 $(-0.053, 0.295)$
lag COH	0.023	0.023 $(-0.054, 0.105)$	0.027 $(-0.045, 0.105)$
NO_3	0.072	0.072 $(-0.046, 0.205)$	0.062 $(-0.047, 0.182)$
SO_4	0.058	0.058 $(-0.049, 0.179)$	0.067 $(-0.027, 0.170)$
o3ppbgt60		0.018 $(-0.183, 0.269)$	0.024 $(-0.141, 0.220)$
Summer[e]			
PM_{10}	0.100	0.100 $(-0.161, 0.444)$	0.088 $(-0.170, 0.427)$
$PM_{2.5}$	-0.077	-0.076 $(-0.326, 0.266)$	-0.087 $(-0.453, 0.524)$
pm2.5aug		0.192 $(-0.172, 0.716)$	0.173 $(-0.202, 0.723)$
lag COH	0.128	0.128 $(-0.085, 0.391)$	0.116 $(-0.094, 0.376)$
NO_3	0.316	0.316 $(0.000, 0.731)$	0.296 $(-0.019, 0.712)$
SO_4	0.107	0.108 $(0.002, 0.224)$	0.108 $(-0.009, 0.238)$
o3ppbgt60	0.107	0.074 $(0.015, 0.137)$	0.094 $(0.022, 0.171)$
Fall[f]			
PM_{10}	0.071	0.071 $(-0.096, 0.269)$	0.095 $(-0.074, 0.295)$
$PM_{2.5}$	0.039	0.039 $(-0.191, 0.335)$	0.043 $(-0.250, 0.448)$
pm2.5aug		0.152 $(-0.052, 0.399)$	0.191 $(-0.027, 0.459)$
lag COH	0.082	0.081 $(-0.018, 0.191)$	0.077 $(-0.021, 0.185)$
NO_3	-0.131	-0.131 $(-0.269, 0.033)$	-0.099 $(-0.239, 0.067)$
SO_4	0.030	0.030 $(-0.058, 0.126)$	0.036 $(-0.060, 0.142)$
o3ppbgt60		0.036 $(-0.031, 0.108)$	0.038 $(-0.029, 0.109)$
Winter[g]			
PM_{10}	0.065	0.065 $(-0.004, 0.138)$	0.057 $(-0.006, 0.124)$
$PM_{2.5}$	0.046	0.046 $(-0.054, 0.155)$	0.034 $(-0.057, 0.135)$
pm2.5aug		0.075 $(0.000, 0.154)$	0.064 $(-0.004, 0.137)$
lag COH	0.036	0.036 $(-0.001, 0.074)$	0.036 $(0.004, 0.070)$
NO_3	0.073	0.073 $(0.006, 0.145)$	0.062 $(0.002, 0.126)$
SO_4	-0.004	-0.004 $(-0.088, 0.088)$	-0.011 $(-0.101, 0.088)$
o3ppbgt60		-0.299 $(-0.892, 3.534)$	-0.337 $(-0.917, 4.301)$

[a] The confidence intervals are c ± 2*s, where for GAM, s = abs(c)/sqrt(F), the F obtained from applying ANOVA to the models with and without the pollutant variable, and for GLM, s = standard error provided by S-Plus.

[b] Relative risks calculated by exp(b*Δp) − 1, where b is the pollutant coefficient from the GAM or GLM regression, and Δp = 50 for PM_{10} and 50 * sd(p)/sd(pm10) for other pollutants, p. For example, sd(pm2.5) = 13, sd(pm10) = 23, so for pm2.5, Δp = 50 * 13/23 = 28. Single asterisks indicate statistical significance at the 0.05 level, double asterisks at the 0.01 level.

[c] All models include a 7 *df* smoothing spline for trend, 12 *df* smoothing spline for season, 3 *df* smoothing spline for minimum temperature, and 2 *df* smoothing spline for maximum temperature.

[d] February, March, April.

[e] May, June, July.

[f] August, September, October.

[g] November, December, January.

Table 4. Comparison of *P* Values for Models Evaluating Daily Mortality from Nonaccidental Causes with Same-Day Pollutant Alone Derived from Different Methods

	t values[a]	F test[b]	Deviance[c]	Simulation[d]	Confidence Limit Lower 95%	Upper 95%	Simulation Size
PM_{10}	0.0017	0.0017	0.0016	0.0015	0.0005	0.0039	2000
$PM_{2.5}$	0.0105	0.0122	0.0145	0.0063	0.0046	0.0087	6000
pm2.5aug	0.0016	0.0017	0.0015	0.0015	0.0005	0.0039	2000
COH	0.1557	0.1786	0.1729	0.1380	0.1180	0.1610	1000
NO_3	0.0016	0.0022	0.0021	0.0020	0.0012	0.0035	6000
SO_4	0.0195	0.0207	0.0196	0.0210	0.0014	0.0320	1000
NO_2	0.0391	0.0599	0.0571	0.0417	0.0351	0.0495	3000
CO	0.1127	0.1410	0.1365	0.1130	0.0990	0.1280	2000
O_3	0.0389	0.0654	0.0624	0.0390	0.0314	0.0485	2000
$PM_{10-2.5}$	0.4912	0.4682	0.4783	0.4990	0.4680	0.5300	1000
o3ppbgt60	0.0005	0.0013	0.0012	0.0008	0.0003	0.0022	4000

[a] Ratio of estimated coefficient to standard error estimate from Table A3 in Fairley 1999.

[b] F test from ANOVA comparing GAMs with and without the pollutant variable.

[c] Assuming that the deviance has a chi-squared distribution with 1 *df*.

[d] Simulations of the null distribution of the pollutant coefficient. Confidence limits based on the binomial distribution.

same-day, single pollutant variables (the same-day, New GAM column of Table 1a). Also included are *P* values using the simple assumption that the change in deviance has a chi-squared distribution with 1 *df*. In addition, the simulation results are shown along with simulation size and confidence intervals based on the binomial.

Table 4 shows that the *t* tests gave results similar to the simulations. The F tests and chi-squared tests gave effectively the same results but appear conservative. In the case of same-day NO_2 and oz8hr, this small difference between the two tests meant the difference between significance and nonsignificance. In the case of the $PM_{2.5}$ coefficient, all the tests appear conservative: the upper confidence bound for the *P* value from the simulation is 0.0087, less than any of the other *P* value estimates. The more extensive simulation, where the entire model-fitting process was simulated, resulted in 8 of 1000 runs in which the simulated $PM_{2.5}$ coefficient was greater than the observed $PM_{2.5}$ coefficient in absolute value. The extensive simulation also suggests that the $PM_{2.5}$ *P* value is, if anything, overestimated using the F test, deviance or *t* value approaches. In the case of the pm2.5aug coefficient, all 4 methods gave similar *P* values, between 0.0015 and 0.0017. The more extensive simulation yielded 2 values out of 1100 greater in absolute value than the observed for an approximate *P* value of 0.0018, very similar to the other results.

SENSITIVITY ANALYSIS: USE OF SINGLE TREND/SEASON SMOOTH

An analysis was performed to determine the effect of using a single smooth for season/trend instead of two. Again, smoothing splines were used and AIC was used as the stopping rule. The minimum AIC of 3104.6 was achieved with 75 *df*, compared with an AIC of 3079.8 using 7 *df* for trend and 12 *df* for season. When temperature terms and $PM_{2.5}$ were added to this model, the estimated $PM_{2.5}$ RR increase dropped from 0.092 to 0.071 and was no longer statistically significant. The estimated RR for pm2.5aug dropped from 0.083 to 0.065 but remained statistically significant.

Figure 2 shows the crude partial ACFs for nonaccidental mortality, the residuals from the 7 *df* trend/12 *df* season fit, and the residuals from the 75 *df* trend/season fit, respectively. Figure 2A shows highly significant autocorrelation in the daily nonaccidental mortality data, with a first order coefficient of about 0.20. Figure 2B shows a first order ACF of just over 0.04, borderline statistically significant, but otherwise an apparently random pattern of positive and negative ACF. Figure 2C shows almost all negative coefficients. The 7 *df* trend/12 *df* season fit had a deviance of 3038.5, similar to the deviance from fitting a single series with 37 *df* (deviance 3039.4), close to the 4 *df* per year suggested by Joel Schwartz at the recent US Environmental Protection Agency (EPA) Workshop on GAM-
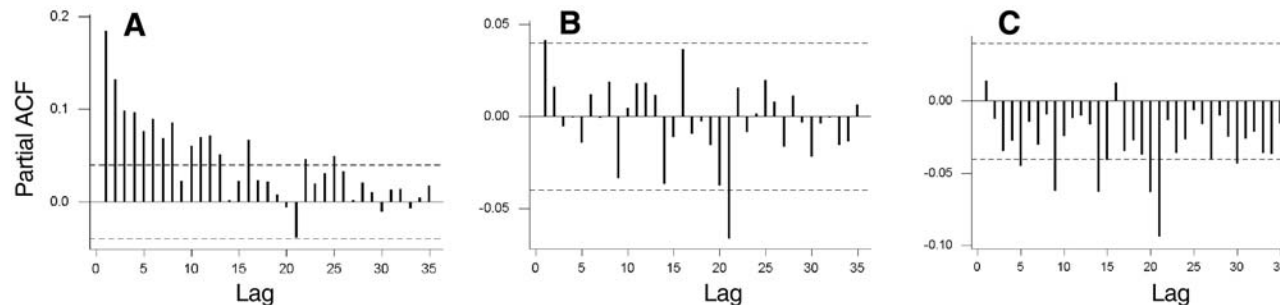
Figure 2. Partial autocorrelation coefficients (partial ACFs). A. Partial ACF of nonaccidental mortality. **B.** Partial ACF of residual *df*: 12 season, 7 time. **C.** Partial ACF of residuals from fitting time with 75 *df*.

Related Statistical Issues in PM Epidemiology, November 4–6, 2002 (Research Triangle Park NC).

DISCUSSION

The basic conclusions of the previous analysis were unchanged: Considered individually, every criteria pollutant (either same-day or lagged by one day) was significantly related to daily mortality. $PM_{2.5}$ and NO_3 continued to be significantly related to mortality in conjunction with the other pollutants.

The fact that the GAM results were virtually identical to the GLM results underscores the conclusion reached in the EPA Particulate Matter Criteria Document (1996), which considered a range of studies analyzing the effects of model choice in short-term mortality studies:

> Differences in model specification may produce important differences in estimates of effects. The general concordance of PM effects estimates, particularly in the analysis of short-term mortality studies, is a consequence of certain appropriate choices in modeling strategy that most investigators have adopted using several different types of standardized models (GLM, LOESS, etc.).

The inferences used in this analysis may be conservative. Although there is no gold standard, the simulations should be the closest to a gold standard because they do not rely on an asymptotic distribution. The only limitation is that the simulations did not account for the uncertainty in the underlying parameters (ie, it did not include simulation of fitting the null model). However, the more extensive simulation of the model-building process of selecting the *df* for trend, season and weather using AIC gave similar results to the simpler simulation. Using the ANOVA feature of S-Plus or assuming the change in the deviance has a chi-square distribution that

gave more conservative results than those from the simulations. Inferences using the GLM-based standard errors agreed closely with the simulations in most cases, suggesting that they do not always produce conservative results. This finding suggests that the GLM-based standard errors may be acceptable for drawing inferences. Perhaps these could be incorporated as defaults or options in statistics packages that offer GAM.

The one modeling choice that does make a substantial difference is using two time smooths—one for trend and one for season—rather than a single time smooth. Using a single time smooth results in considerably lower PM coefficient estimates and lower statistical significance. However, the single time smooth model that minimized the AIC had 75 *df*, which is more than 9 per year from 1989 to 1996. This, according to the discussion at the GAM workshop, may constitute overfitting. In fact, Figure 2C shows evidence of overfitting in that almost all the autocorrelation coefficients are negative. Also, the AIC for the single time-term model is 3104.6, considerably poorer than the AIC for the 7 *df* trend/12 *df* season model, 3079.8.

Parsimony is an established scientific rule. By Ockham's Razor, if there is a choice between similar-fitting models, the one with fewer terms should be favored. And a previously suspected cause (in this case PM) should be favored over an unknown cause (for which time is a surrogate). In practical terms, the question to address is how much of the short-term mortality variation to attribute to PM vis-à-vis some other covariate represented by additional degrees of freedom in the time smooth. The fitting process begins with a fit of time/season only. This tends to add more degrees of freedom for these terms because every day is included, not the 1 in 6 days where $PM_{2.5}$ data are available. Secondly, AIC is a liberal criterion, which includes variables that are not statistically significant. Thus, the modeling approach allows a reasonable chance for the other covariate to demonstrate its existence. Yet pm2.5aug remains statistically

significant, even in the 75 *df* trend/season model, and the estimated PM effect drops by about a quarter, rather than disappearing entirely. Thus, the two-smooths model with fewer parameters and higher PM coefficients appears preferable according to the Ockham's Razor.

The new ozone variable, o3ppbgt60, was found to be statistically significant even in a regression jointly with $PM_{2.5}$ or NO_3. Two other thresholds were tried—40 ppb and 80 ppb—both of which produced statistically significant results in one-pollutant models but somewhat smaller RRs. Although o3ppbgt60 is quite skewed, the results do not appear to be a function of a few outliers; a regression using log(o3ppbgt60) also yielded statistically significant results.

REFERENCES*

Chambers JM and Hastie TJ. Statistical Models in S. Pacific Grove: Wadsworth & Brooks/Cole. (Chapter 7.4.4, page 303)

Cox DR and Hinkley DV 1974. Theoretical statistics. London: Chapman and Hall.

Dominici F, Zeger SL, and Samet JM. 2000. Combining evidences on air pollution and daily mortality from the largest 20 US cities: A hierarchical modeling strategy. Royal Statistical Society, Series A. 163:263-302

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On the use of Generalized Additive Models in time series of air pollution and health. Am J Epidemiol 156:193-203

Environmental Protection Agency (US). 1996. Air quality criteria for particulate matter, Vol 3.pp 13-53:13-54, EPA/600/P-95/001cF, April 1996.

Fairley D. 1999. Daily mortality and air pollution in Santa Clara County, California: 1989–1996. Environ Health Perspect 107:637–641.

Kelsall JE, Samet JM, Zeger SL, and Xu J. Air pollution and daily mortality in Philadelphia, 1974-1988. Am J Epidemiol 146:750-762.

Moolgavkar S. 2000. Air pollution and daily mortality in three U.S. counties. Environ Health Perspect 108:777-784.

Saez M, Ballester F, Barelo MA, Perez-Hoyos S, Bellido J, Tenias JM, Ocaña R, Figueiras A, Arribas F, Aragones N, Tobias A, Cirera L, Cañada A. "A combined analysis of the short-term effects of photochemical air pollutants on mortality within the EMECAM Project," Environ Health Perspect. 110:221-228 (2002).

Samet JM, Dominici F, Curriero F, Coursac I, and Zeger SL. 2000. Particulate air pollution and mortality: Findings from 20 U.S. cities. New England Journal of Medicine, 343:1742-1757.

Schwartz J. 1994. Nonparametric smoothing in the analysis of air pollution and respiratory illness. Can J Stat 22:471-488.

ABBREVIATIONS AND OTHER TERMS

ACF	autocorrelation coefficient
AIC	Akaike information criterion
ANOVA	analysis of variance
CL	confidence limit
CO	carbon monoxide
COH	coefficient of haze
df	degree of freedom
EPA	Environmental Protection Agency (US)
GAM	generalized additive model
GLM	generalized linear model
ICD	*International Classification of Diseases*
NO_2	nitrogen dioxide
NO_3	nitrate
ns	natural spline
O_3	ozone
o3ppbgt60	ozone ppb-hours greater than 60 ppb summed for each day
oz8hr	8-hour maximum ozone
PM	particulate matter
PM_{10}	particulate matter < 10 μm in diameter
$PM_{2.5}$	particulate matter < 2.5 μm in diameter
pm2.5aug	$PM_{2.5}$ measurements augmented with $PM_{2.5}$ predicted from COH and PM_{10}
ppb	parts per billion
R^2	multivariate coefficient of variation
RR	relative risk
SO_4	sulfate

---

* Bold type identifies publication containing the original analyses revised in this short communication report.

Ambient Pollution and Reduced Heart Rate Variability

Diane R Gold, Joel Schwartz, Augusto Litonjua, Richard Verrier, and Antonella Zanobetti

ABSTRACT

We investigated associations between ambient pollution levels and heart rate variability (HRV*) in a repeated measures study of elderly Boston residents and published our study results in *Circulation*[1]. The study protocol involved 25 minutes per week of continuous Holter electrocardiogram (ECG) monitoring including 5 minutes of rest, 5 minutes of standing, 5 minutes of exercise outdoors, 5 minutes of recovery, and 20 cycles of slow breathing. Measures of HRV included the standard deviation of normal intervals between adjacent R waves on an ECG (RR intervals) (SDNN) and the square root of the mean of the squared differences between adjacent normal RR intervals (r-MSSD). Ambient particle and ozone increases were associated with decreased HRV, suggesting pollution-related autonomic imbalance. In the published article we presented results from random effects analyses that did not utilize generalized additive models (GAMs), demonstrating our ability to replicate our main findings derived from GAMs. To address any further concerns regarding secondary results from GAMs that might use inadequate default convergence criteria and/or have variance estimation issues, we performed a comprehensive revised analysis of our study and found no significant differences in the estimates of the effects of ambient pollution (airborne particulate matter with an aerodynamic diameter equal to or less than 2.5 μm [$PM_{2.5}$] or ozone) on HRV.

INTRODUCTION

Reduced HRV is a predictor of increased risk for cardiovascular mortality and morbidity[2,3]. In a recently published study[1], we assessed the relation between short-term changes in multiple air pollutants and short-term changes in HRV in a community-based elderly population. Ambient particle and ozone increases were associated with decreased HRV, suggesting pollution-related autonomic imbalance.

METHODS

STUDY SAMPLE AND PROTOCOL

In this paper we present a revised analysis of data first presented in published form in *Circulation*[1]. Significant portions of this section were first presented in that article. The methods were as follows. Between May and July 1997 volunteers were recruited from a Boston housing community, where the screening and testing office were located, on the ground floor of an apartment building. A questionnaire was administered regarding medications, pulmonary and cardiac symptoms, and smoking history. A resting 12-lead ECG was performed. Exclusion criteria included unstable angina, atrial flutter, atrial fibrillation, paced rhythm, and left bundle branch block. Inclusion criteria included the ability to walk on level ground. Of the 31 individuals who were screened, 21 entered the repeated measures study.

Each participant was given a day of the week and time when weekly testing would be performed. Subjects were tested June through September from 9:00 am to 2:00 pm, Monday through Friday, by a team of two technicians or physicians. Participants were administered a brief questionnaire regarding chest pain, doctor's visits, hospital visits, medication changes, and whether medication had been taken that morning. Continuous Holter monitoring with electrodes in a modified V5 and AVF position was performed during a protocol: (1) <u>Five minutes of rest.</u> Respiratory rate and three supine blood pressures were measured using a mercury column sphygmomanometer; (2) <u>Five minutes of standing.</u> After two minutes of equilibration, standing blood pressure was measured three times; (3) <u>Five minutes of exercise outdoors.</u> If the participant felt able, a

---

standard walk was performed, involving one climb up a slight incline; (4) <u>Five minutes of recovery.</u> The participant lay down again and respiratory rate was recorded; (5) <u>Three minutes and twenty seconds of slow breathing.</u>[4] During each of 20 respiratory cycles, the participant was asked to breathe in for five seconds and then out for five seconds, with technician coaching. The slow-breathing portion enabled us to evaluate whether the effects of pollution on HRV were independent of respiratory rate, which might also be influenced by pollution levels.

PROCESSING OF HOLTER RECORDINGS

Using a Marquette MARS Workstation, a trained engineer reviewed and, when necessary, corrected automatically determined readings of QRS complexes. Regions of noise and artifact (< 1% of data) were eliminated. After correction, software facilities on the MARS were used to export beat timing and annotation information for analysis and creation of outcome variables through custom PC-based software written in the C language. Only normal-to-normal (NN) intervals between 150 and 5000 milliseconds with NN ratios between 0.8 and 1.2 were included for analysis of HRV. No tape contained more than 1% premature beats.

Two time-domain measures of HRV were obtained. The SDNN and r-MSSD were calculated from all normal RR intervals for each portion of the protocol and the overall protocol.

EXPOSURE MONITORING

Airborne particulate matter with an aerodynamic diameter equal to or less than 10 μm (PM_{10}) and $PM_{2.5}$ were measured continuously 6 km from the study site using a model 1400A tapered element oscillating microbalance (TEOM). Since the TEOM sample filter is heated to 50℃, a season-specific correction was used to compensate for the loss of semivolatile mass that occurs at this temperature.[5] Calibration factors were obtained by regressing continuous $PM_{2.5}$ and PM_{10} concentrations (averaged over 24 hours) on the corresponding collocated integrated 24-hour Harvard impactor low-volume Teflon-filter gravimetric measurements:

corrected $PM_{2.5}$ = (measured $PM_{2.5}$ + 2.00)/0.944
for May through August (r^2 = 0.99).

Coarse matter was calculated by subtracting $PM_{2.5}$ from PM_{10}. Continuous carbon monoxide data were collected within a quarter of a mile of the participant residence with a ThermoEnvironmental (Franklin MA) model 48 gas analyzer using a US Environmental Protection Agency (EPA) reference method. Ozone, nitrogen dioxide (NO_2), sulfur dioxide (SO_2), temperature and relative humidity measurements were obtained from the Massachusetts Department of Environmental Protection local monitoring site 4.8 miles from the study site.

STATISTICAL ANALYSIS

Continuous or categorical predictor variables assumed to be time invariate (or changing slowly) included age, sex, race/ethnicity, body mass index, and the diagnostic categories derived from answers to the screening questionnaire and summarized in Table 1 of the original article[1]. Time-varying predictors included air pollutants, temperature, relative humidity, and medication use.

Data analyses dealt with the variable number of repeated measures for each subject. Whereas individual covariates were available for each subject and were used to control for individual differences, measured covariates were thought unlikely to explain all interindividual differences. Each subject was not seen on each day of the study, but rather once or twice a week, thus creating the potential for variations in persons to be seen each day that could confound time-varying exposures such as air pollution. We controlled for these variations primarily by constructing fixed-effect models,[6] fitting an individual intercept for each subject while still adjusting for time-varying covariates and individual traits (the most important of which was medication use).

Fixed-effects models have the advantage of adjusting for both measured and unmeasured time-invariant characteristics of the individual but also have the disadvantage of not providing estimates for specific measured time-invariant subject characteristics. We used the SAS mixed procedure[7] to construct a second set of random-effects models to evaluate the sensitivity of air-pollution results to the choice of model and to define the primary effects and interactions with air pollution of subject characteristics, the individual effects of which could not be evaluated in a fixed-effect model. Because multiple measurements were taken for each subject, and because those measurements may not be independent, a random-subject effect was used in these regression analyses.

Weather and air pollution are continuous-exposure measures and may not be linearly related to electrophysiologic measures. To test this assumption, we repeated the analyses using GAMs in S-Plus.[6] A GAM fits the outcome as a sum of functions of each predictor that are not required to be linear. The shape of these functions is estimated from the data using nonparametric smoothing, and the significance of any deviations from linearity can be tested using nonparametric F tests. Fixed-effects models used nonparametric smoothing to adjust for temperature, since temperature did not always relate linearly with heart rate and HRV.

In the original article[1], GAM modeling with default convergence criteria and nonparametric smoothing to adjust for temperature was used in analyses presented in Table 4 ("Ambient $PM_{2.5}$ as a Predictor of Heart Rate and of Heart Rate Variability for 5 Protocol Periods") and Table 6 ("Estimated Effects of Pollution on Heart Rate Variability in Single and Multiple Pollutant Models"). In Table 5 ["Predictors of Heart Rate and Heart Rate Variability (r-MSSD), Random Effects Models"] from the original article, we demonstrated our ability to replicate our main findings derived from GAMs in our random-effects models that utilized neither GAM nor nonparametric smoothing. Elevated $PM_{2.5}$ levels during the hour of testing and during the 3 hours before testing (4-hour $PM_{2.5}$) were associated with reduced HRV. Both in fixed modeling (original Tables 4 and 6) and in random effects modeling (original Table 5) (using a smooth vs linear term for temperature), we estimated a 4-millisecond reduction in r-MSSD during the first rest period that resulted in an interquartile difference in 4-hour $PM_{2.5}$ of 14 µg/m^3.

To address any further concerns regarding secondary results from GAMs that might use inadequate default convergence criteria or have variance estimation issues, we performed a comprehensive reanalysis of our study and found no significant differences in the estimates of the effects of ambient pollution ($PM_{2.5}$ or ozone) on HRV. First, to address the question of sensitivity of the effect size estimates to the choice of convergence criteria and maximum iterations in GAMs, we refit the originally reported fixed-effects models, changing only the convergence criteria. We used the stricter convergence criteria and number of iterations recommended by Dominici and coworkers[8] (eg, 10^{-15} and 1000, respectively).

In some models we had used a locally weighted running-line smoother (LOESS)[9] to model temperature. To address the question of sensitivity to the incorrect estimation of standard errors, we repeated the analyses using both the stricter convergence criteria and, where temperature was included in the model, natural spline (with 3 degrees of freedom [*df*]) was used instead of LOESS to adjust for temperature.

RESULTS AND CONCLUSION

The original results published in *Circulation*[1] are presented here in Tables 1 and 2. Table 1 presents the original and revised results from Table 4 in the original article; Table 2 presents the original and revised results from Table 6 in the original article. In revised analyses not including temperature, results were essentially unchanged using the stricter convergence criteria (Table 1, models 1

and 2). Original analyses including temperature are shown in Table 1 (only model 3) and Table 2. In revised analyses for models that included temperature, when we used the stricter convergence criteria but still adjusted for temperature with LOESS (nonparametric smoothing), results were essentially unchanged from the original results (results not shown). We present the results of revised analyses that include temperature in which we use both the stricter convergence criteria and natural spline to adjust for temperature (Table 1, model 3; Table 2). The magnitude and significance of the association of $PM_{2.5}$ and ozone with reduced HRV (r-MSSD) was essentially unchanged by the use of both stricter convergence criteria and natural spline rather than LOESS for temperature. We conclude that the use of stricter convergence criteria and natural spline significantly influenced neither the magnitude nor the precision of the estimate of the association of ambient pollution ($PM_{2.5}$ or ozone) with reduced HRV.

REFERENCES*

1. **Gold DR, Litonjua A, Schwartz J, Lovett E, Larson A, Nearing B, Allen G, Verrier M, Cherry R, Verrier R. 2000. Ambient pollution and heart rate variability. Circulation 101:1267–1273.**

2. Tsuji H, Venditti FJ, Jr., Manders ES, et al. Reduced heart rate variability and mortality risk in an elderly cohort. The Framingham Heart Study. Circulation 1994; 90:878-83.

3. Tsuji H, Larson MG, Venditti FJ, Jr., et al. Impact of reduced heart rate variability on risk for cardiac events. The Framingham Heart Study. Circulation 1996; 94:2850-5.

4. Freeman R, Weiss ST, Roberts M, Zbikowski SM, Sparrow D. The relationship between heart rate

---

* Bold text identifies publication containing the original analyses revised in this short communication report.

Table 1. Ambient $PM_{2.5}$ as a Predictor of Heart Rate and Heart Rate Variability for 5 Protocol Periods[a]

Outcome Variable	Predictor Variable ($\mu g/m^3$)	Model 1		Model 2 (with heart rate)		Model 3 (with 24-hr temperature)	
		Original Coefficient (SE)	Revised Coefficient (SE)	Original Coefficient (SE)	Revised Coefficient (SE)	Original Coefficient (SE)	Revised Coefficient (SE)
Heart rate (beats/min)	24-hr $PM_{2.5}$						
First rest		−0.10 (0.04)	−0.10 (0.04)			−0.15 (0.06)	−0.15 (0.06)
Standing		−0.15 (0.05)	−0.15 (0.05)			−0.20 (0.07)	−0.19 (0.08)
Exercise		−0.10 (0.05)	−0.10 (0.05)			−0.21 (0.06)	−0.22 (0.07)
Second rest		−0.02 (0.05)	−0.02 (0.05)			−0.11 (0.07)	−0.11 (0.07)
Slow breathing		−0.04 (0.04)	−0.04 (0.04)			−0.09 (0.05)	−0.10 (0.06)
Overall		−0.09 (0.04)	−0.09 (0.04)			−0.16 (0.06)	−0.16 (0.06)
SDNN (ms)[b]	4-hr $PM_{2.5}$						
First rest		−0.28 (0.16)	−0.28 (0.16)	−0.22 (0.16)	−0.22 (0.16)	−0.25 (0.18)	−0.25 (0.18)
Standing		−0.09 (0.15)	−0.09 (0.15)	−0.17 (0.13)	−0.17 (0.13)	0.004 (0.17)	0.002 (0.17)
Exercise		−0.25 (0.15)	−0.25 (0.15)	−0.27 (0.14)	−0.27 (0.14)	−0.24 (0.17)	−0.25 (0.17)
Second rest		−0.29 (0.20)	−0.29 (0.20)	−0.27 (0.20)	−0.27 (0.20)	−0.11 (0.22)	−0.11 (0.22)
Slow breathing		−0.16 (0.14)	−0.16 (0.14)	−0.16 (0.14)	−0.16 (0.14)	−0.20 (0.16)	−0.21 (0.16)
Overall[c]		−0.24 (0.10)	−0.24 (0.10)	−0.26 (0.09)	−0.26 (0.09)	−0.17 (0.11)	−0.18 (0.11)
r-MSSD (ms)[b]	4-hr $PM_{2.5}$						
First rest		−0.28 (0.09)	−0.28 (0.09)	−0.27 (0.09)	−0.27 (0.09)	−0.28 (0.11)	−0.27 (0.11)
Standing		−0.28 (0.11)	−0.28 (0.11)	−0.32 (0.11)	−0.32 (0.11)	−0.29 (0.12)	−0.30 (0.13)
Exercise		−0.09 (0.08)	−0.09 (0.08)	−0.10 (0.08)	−0.10 (0.08)	−0.13 (0.09)	−0.14 (0.09)
Second rest		−0.29 (0.12)	−0.29 (0.12)	−0.27 (0.12)	−0.27 (0.12)	−0.25 (0.14)	−0.25 (0.14)
Slow breathing		−0.35 (0.13)	−0.35 (0.13)	−0.35 (0.14)	−0.35 (0.14)	−0.42 (0.15)	−0.42 (0.15)
Overall[c]		−0.25 (0.08)	−0.25 (0.08)	−0.26 (0.08)	−0.26 (0.08)	−0.29 (0.08)	−0.29 (0.08)

[a] All repeated-measures regression models contained fixed-effects indicator variables for each of 21 participants and variables for whether the participant took a β blocker, calcium channel blocker, angiotensin converting enzyme inhibitor, or sympathomimetic medication on the testing day. All revised results are from models using the convergence criteria and number of iterations recommended by Dominici and colleagues[8] (10^{-15} and 1000, respectively). No revised results used LOESS. Where temperature was included (model 3), revised results including temperature used natural spline with 3 *df*.

[b] Mean of the estimates of all 5 portions of the protocol, weighted by the standard errors.

[c] SDNN = standard deviation of normal RR intervals; ms = millisecond; r-MSSD = square root of the mean of the squared differences between adjacent normal RR intervals.

Table 2. Estimated Effects of Pollution on Heart Rate Variability in Single and Multiple Pollutant Models[a]

Outcome Variable (mean value)	Model	Predictor Variable	Original Estimated Effect (SE)	Original P Value	Revised Estimated Effect (SE)	Revised P Value
Heart rate, first rest period (beats/min)	1	24-hr $PM_{2.5}$	−1.8 (0.7)	0.01	−1.8 (0.8)	0.02
	2	24-hr NO_2	−1.4 (0.6)	0.02	−1.4 (0.7)	0.04
	3	24-hr SO_2	−1.0 (0.5)	0.03	−1.0 (0.5)	0.03
	4	24-hr $PM_{2.5}$	−1.6 (0.7)	0.03	−1.6 (0.8)	0.04
		24-hr NO_2	−1.0 (0.6)	0.09	−1.0 (0.7)	0.13
	5	24-hr $PM_{2.5}$	−1.6 (0.7)	0.03	−1.6 (0.8)	0.04
		24-hr SO_2	−0.8 (0.5)	0.09	−0.8 (0.5)	0.10
Overall heart rate (beats/min)	1	24-hr $PM_{2.5}$	−1.9 (0.7)	0.01	−1.9 (0.7)	0.01
	2	24-hr NO_2	−1.0 (0.6)	0.10	−0.9 (0.7)	0.20
	3	24-hr SO_2	−0.5 (0.5)	0.30	−0.5 (0.5)	0.31
	4	24-hr $PM_{2.5}$	−1.8 (0.7)	0.01	−1.9 (0.8)	0.01
		24-hr NO_2	−0.6 (0.6)	0.32	−0.5 (0.7)	0.49
	5	24-hr $PM_{2.5}$	−1.9 (0.7)	0.01	−2.0 (0.8)	0.01
		24-hr SO_2	−0.2 (0.5)	0.60	−0.2 (0.5)	0.63
r-MSSD, first rest period (ms)[b]	1	4-hr $PM_{2.5}$	−4.0 (1.5)	0.009	−3.9 (1.5)	0.01
	2	1-hr O_3	−3.0 (1.8)	0.11	−3.0 (1.9)	0.12
	3	4-hr $PM_{2.5}$	−3.7 (1.6)	0.02	−3.7 (1.6)	0.03
		1-hr O_3	−2.9 (1.7)	0.10	−2.9 (1.8)	0.12
r-MSSD, slow breathing period (ms)[b]	1	4-hr $PM_{2.5}$	−6.1 (2.2)	0.006	−6.0 (2.2)	0.007
	2	1-hr O_3	−5.9 (2.3)	0.01	−5.8 (2.4)	0.02
	3	4-hr $PM_{2.5}$	−5.4 (2.2)	0.02	−5.4 (2.2)	0.02
		1-hr O_3	−5.5 (2.4)	0.03	−5.4 (2.5)	0.03

[a] All repeated measures regression models contained fixed effects indicator variables for each of 21 participants and variables for whether the participant took a β blocker, calcium channel blocker, angiotensin converting enzyme inhibitor, or sympathomimetic medication on the testing day and for 24-hr temperature. All original results used nonparametric smoothing to adjust for temperature. No revised results used LOESS. All revised results (1) used stricter convergence criteria and the number of iterations recommended by Dominici and colleagues[8] (eg, 10^{-15} and 1000, respectively) and (2) included temperature using natural spline with 3 *df* (*not* LOESS). Effect was estimated for interquartile changes in the ambient pollutant. The interquartile range (Q^1–Q^3) for 24-hr $PM_{2.5}$ was 12 μg/m³ and for 4-hr $PM_{2.5}$ was 14.35 μg/m³.

[b] r-MSSD = square root of the mean of the squared differences between adjacent normal RR intervals; ms = millisecond.

variability and measures of body habitus. Clin Auton Res 1995; 5:261-6.

5. Allen G, Sioutas C, Koutrakis P, Reiss R, Lurmann FW, Roberts PT. Evaluation of the TEOM method for measurement of ambient particulate mass in urban areas. J Air Waste Manage Assoc 1997; 47:682-689.

6. Statistical Sciences, S-Plus User's Manual. Version 3.2. Seattle: StatSci, MathSoft, 1993.

7. SAS Institute. SAS/STAT software: changes and enhancements through release 6.12. SAS Institute, Cary, NC. 1997.

8. Dominici F, McDermott A, Zeger SL, Samet JM. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiol 2002; 156:193-203.

9. Cleveland WS, Devlin SJ. Robust locally-weighted regression and smoothing scatter plots. J Am Stat Assoc 1988; 74:829-36.

ABBREVIATIONS AND OTHER TERMS

df	degree of freedom
ECG	electrocardiogram
EPA	Environmental Protection Agency (US)
GAM	generalized additive model
HRV	heart rate variability
LOESS	locally weighted running-line smoother
NN	normal-to-normal
NO_2	nitrogen dioxide
O_3	ozone
$PM_{2.5}$	airborne particulate matter with an aerodynamic diameter equal to or less than 2.5 µm
PM_{10}	airborne particulate matter with an aerodynamic diameter equal to or less than 10 µm
r^2	coefficient of variation for bivariate analysis
r-MSSD	square root of the mean of the squared differences between adjacent normal RR intervals
RR interval	interval between adjacent R waves on an ECG
SDNN	standard deviation of normal RR intervals
SO_2	sulfur dioxide
TEOM	tapered element oscillating microbalance

Revised Analysis of the Montreal Time-Series Study

Mark S Goldberg and Richard T Burnett

ABSTRACT

Methods: We conducted a revised analysis of the association between particulate mass and nonaccidental mortality over 10 years, 1984 to 1993, in Montreal, Quebec. The original publication results were based on nonparametric locally weighted smoothers (LOESS*) for time and weather variables using the S-Plus default convergence criteria. In the reanalysis, these were replaced with natural spline (parametric) models. The primary reanalysis made use of the original criteria for selecting models (temporal filter with the minimum Bartlett test for white noise) and, as before, the weather variables were selected from those that produced the minimum Akaike information criterion (AIC). These parametric log-linear Poisson models accounted for day of the week, calendar year, and overdispersion. A series of sensitivity analyses were conducted using different degrees of freedom (*df*) for the natural spline temporal filter and different forms of the weather variables.

Results: The associations for mortality were not very sensitive to the amount of temporal smoothing but were highly sensitive to the functional form used to adjust for weather. Most of the originally reported associations, except for congestive heart failure, were highly attenuated when a natural spline for weather (usually mean daily temperature) was used. The response curve for mean temperature was rather flat until about 15°C, at which point it increased dramatically. Other models, which assumed a threshold effect of 22°C or 25°C, did not show such dramatic decreases in the air pollution effect.

Conclusions: Weather had such a profound effect on the particle associations for several possible reasons: (1) the results represent the true effect for temperature that was

missed in our original generalized additive model (GAM) analysis because concurvity among covariates was not detected; (2) temperature affects mortality but has a threshold response function (effectively zero below reasonably high values of near 22°C or 25°C), so the natural spline models overfit the data and hence remove the air pollution effect; and (3) 1 and 2 are correct to varying degrees, but the observed confounding effects were accentuated because of transference of causal effects from less-precisely to more-precisely measured variables. Prior knowledge of the effects of weather on mortality and other health outcomes would greatly benefit the development of strategies for selecting the appropriate functional form for weather in the presence of modest correlations between covariates. It is also clear that GAMs do not always provide unbiased estimates of effect and should not be used until the GAM backfitting algorithm has been corrected.

INTRODUCTION

This short communication describes the results of a revised analysis of the association between particle mass and nonaccidental mortality in Montreal from January 1, 1984 to December 31, 1993. The original study was divided into two parts, each of which used a different definition of nonaccidental mortality: (1) cause-specific daily mortality and (2) nonaccidental daily mortality, stratified by underlying health conditions subjects had before death (as defined from data from the Quebec Health Insurance Plan [QHIP]). The description of the original analysis and results were published as HEI Research Report 97 (1) and in other journal articles (2,3,4,5), and one new paper is in press regarding congestive heart failure.

METHODS USED IN THE ORIGINAL STUDY

The results in the original publications were based on time-series analyses that provided estimates of the association between daily variations in mortality with daily

---

variations in air particulates, after controlling for seasonal and subseasonal trends in the mortality time series, daily weather, and gaseous pollutants. We used the GAM framework (using LOESS for all continuous covariates) and selected the span for the LOESS temporal filter that minimized for the residuals the Bartlett test for white noise (6). The weather variables were selected among those (mean temperature, dew point temperature, change in temperature from the previous day, change in barometric pressure from the previous day) across lags 0–5 days for which the value of the AIC was minimized (using a span of 50%).

STATISTICAL METHODS

SCOPE OF THE REANALYSIS

For all of the reanalyses presented here, the air pollution, mortality, and QHIP data are identical to those used in the original publications. The revised analysis was guided by the finding of two errors in the implementation of the GAM backfitting algorithm: insufficiently stringent default convergence criteria (7) and an inability to account for nonlinear correlations between independent variables (referred to as *concurvity*) (8). We believe that the implementation of GAM yielded biased estimates of effect and variances when there are sufficiently strong correlations among covariates, as in the Montreal data. Nevertheless, as requested by the US Environmental Protection Agency (EPA), we have reanalyzed our data using the same functional forms of the original GAMs with more strict convergence criteria (`epsilon and bf.epsilon = 10^{-15}`, `maxit and bf.maxit = 1000`).

The scope of the analysis agreed to with HEI was to attempt to reproduce all relevant tables and figures in the report. Because of severe time constraints, it was recognized that judicious selection of reanalyses of tables and graphs in the original report was necessary and, consequently, this reanalysis covers only certain key features of the original HEI Research Report (7). In particular, we did not reanalyze cause-specific mortality except for all nonaccidental mortality; rather, we focused on the analysis of those subgroups (defined using QHIP data [part 2 of the original report]) that showed associations with particles. In addition, we discovered during the reanalysis that our original assumptions concerning confounding by weather were being seriously challenged and we thus decided to undertake a series of comprehensive sensitivity analyses to appreciate how different models for weather affect our findings.

REVISED ANALYSIS METHODS: USING NATURAL SPLINES TO MODEL CONTINUOUS COVARIATES

In place of the LOESS in the original GAM, we used parametric natural spline models within the context of generalized linear models (GLMs) (9) to capture the potentially nonlinear effects of time and weather on daily mortality. The natural spline models are defined by the degrees of freedom, which are given by the number of knots plus two. The knots are placed evenly throughout the distribution of the variable. Natural spline models have been shown to perform well in time-series mortality studies, as compared with LOESS and nonparametric spline smoothers, in terms of lack of bias in the air pollution effect estimate and the associated standard errors.

The amount of smoothing needed when applying a natural spline function equivalent to that of a LOESS function is difficult to define. Both have a concept of degrees of freedom but nonparametric smoothers are defined in a manner that fits the data better than parametric functions, like natural splines, for the same degrees of freedom. The LOESS function was therefore not replaced with a natural spline having the same number of degrees of freedom. Moreover, because the issue of collinearity was not accounted for properly in the GAM backfitting algorithm, we needed to start the modeling anew. Thus, we used the same type of modeling approach as used in the original analyses both to select the degrees of freedom for the natural splines for the temporal filter as well as for the weather variables.

Unlike the original analysis, however, we added day of the week as an indicator variable and, as before, included a term for calendar year if it improved the fit. For each separate time series, we found the degrees of freedom for the natural spline function on time that produced a minimum value of the Bartlett test for white noise. The number of days with information in a time series was defined as the intersection of the number of days of data for mortality and for the pollutant (we had complete data for weather). A different number of degrees of freedom for the temporal smoother could thus be found for the same endpoint if the number of days with information was different; for example, comparing six-day total suspended particles (TSP) data to every-day data for coefficient of haze (COH). We evaluated a wide range of degrees of freedom before selecting the final one. Table 1 shows an example of this approach for all nonaccidental mortality (complete time series); 88 *df* (8.8 per year, shown in bold type) was selected in this case.

To select the relevant weather variables, we conducted a series of univariate analyses using natural splines, evaluated over lags of 0 to 5 days and a maximum of 6 *df*, adjusting only for temporal trends. We perused these models and selected variables that yielded the minimum AIC. We then investigated pairs of variables and selected the model with the minimum AIC. Table 2 shows the selection procedure for the pairs of weather variables used in the analysis for nonaccidental mortality, and Table 3 shows the modeling of the selected weather variables: mean temperature (4 *df*, lag 0) and relative humidity (linear effect, lag 2). Table 4 shows an example of the different models selected for nonaccidental mortality and the different pollutants.

Because we ended up using a larger number of degrees of freedom than most other investigators (perhaps due to the large seasonal cycles in our data; Figures 1–3), we also conducted sensitivity analyses using other degrees of freedom.

REVISED ANALYSIS RESULTS

Nonaccidental Mortality

Table 5 shows the original results, those from stricter convergence criteria used in the GAMs, and those of the revised analysis based on natural splines (based on Table 15 of the original report). (Appendix Table A.1 shows the distribution of environmental pollutants.) The analyses

Table 1. Daily Nonaccidental Mortality: Selection of the Degrees of Freedom for the Natural Spline (ns) Filter, with Day of the Week and Calendar Year Entered into the Model

					Serial Autocorrelation Coefficients					
ns *df*	Dispersion	Bartlett	*P* value	AIC	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
81	1.09	1.271	0.079	4095.1	0.0379	−0.0045	0.0008	−0.0233	−0.0207	−0.0044
86	1.09	1.194	0.116	4093.7	0.0349	−0.0074	−0.0021	−0.0263	−0.0235	−0.0071
91	1.08	1.015	0.254	4076.0	0.0279	−0.0147	−0.0091	−0.0334	−0.0302	−0.0136
96	1.08	1.087	0.188	4065.7	0.0226	−0.0201	−0.0144	−0.0388	−0.0356	−0.0186
101	1.08	1.162	0.134	4071.0	0.0211	−0.0215	−0.0157	−0.0401	−0.0369	−0.0199
80	1.09	1.137	0.15	4071.6	0.0327	−0.0098	−0.0043	−0.0286	−0.0257	−0.0091
82	1.09	1.147	0.14	4077.4	0.0330	−0.0093	−0.0040	−0.0281	−0.0252	−0.0087
83	1.09	1.155	0.14	4080.5	0.0333	−0.0091	−0.0036	−0.0280	−0.0250	−0.0084
84	1.09	1.133	0.15	4080.3	0.0325	−0.0098	−0.0043	−0.0285	−0.0258	−0.0092
85	1.09	1.044	0.23	4068.0	0.0291	−0.0136	−0.0079	−0.0321	−0.0292	−0.0125
86	1.09	1.194	0.12	4093.7	0.0349	−0.0074	−0.0021	−0.0263	−0.0235	−0.0071
87	1.09	1.047	0.22	4072.7	0.0292	−0.0134	−0.0077	−0.0321	−0.0290	−0.0124
88	**1.08**	**1.001**	**0.27**	**4054.5**	**0.0243**	**−0.0184**	**−0.0126**	**−0.0369**	**−0.0338**	**−0.0169**
89	1.08	0.999	0.27	4070.0	0.0273	−0.0153	−0.0095	−0.0338	−0.0309	−0.0141

Table 2. Daily Nonaccidental Mortality: Selection of the Combination of Weather Variables[a]

First Weather Variable	Lag	*df*	Second Weather Variable	Lag	*df*	Dispersion	AIC
Mean temperature	0	4	Mean RH	2	0	1.05	3943.0
Mean temperature	0	4	Change in pressure	0	5	1.05	3943.8
Maximum temperature	0	4	Change in pressure	0	5	1.05	3955.7
Maximum temperature	0	4	Mean RH	2	0	1.05	3959.9
Dew point	0	3	Mean RH	2	0	1.06	3966.1
Dew point	0	3	Change in pressure	0	5	1.06	3972.1
Mean RH	2	0	Change in pressure	0	5	1.07	4023.0

[a] RH = relative humidity.

Table 3. Daily Nonaccidental Mortality: Example of Modeling the Weather Variables[a]

Model	Residual Deviance	Change in Residual Deviance	Change in *df*	Likelihood Ratio Test *P* value	AIC
Base	3840.3	NA[b]	NA[b]	NA[b]	4036.0
+ Mean temperature (lag 0, natural spline with 4 *df*)	3748.4	91.95	4	0.0	3947.9
+ Relative humidity (lag 2, linear effect)	3736.9	8.93	1	0.003	3940.2

[a] Base model includes coefficient of haze as a linear effect, temporal filter using natural spline with 88 *df* and calendar year as a linear effect.

[b] NA = not applicable.

using the stricter convergence criteria produced slightly lower estimates than the original analysis; however, the conclusions would not have been altered had this approach been used. Results using the natural spline approach yielded attenuated and statistically nonsignificant estimates of effect for the five sets of particle metrics presented, and inclusion of day of the week reduced the estimates further. Figures 4 through 6 compare the original analysis and reanalysis for each lag considered and Figures 7 and 8 show the analyses for the warm and cold seasons, respectively. Table 6 shows little effect of using different temporal filters and degrees of freedom for mean temperature on the results.

Figure 9 shows the functional form of mean temperature using 2 *df* and 4 *df* and shows that the effect of temperature increases dramatically starting around 10°C and 15°C, respectively (these temperatures correspond roughly to a maximum daily temperature of between 13–20°C and 16–26°C, respectively; the Pearson correlation between mean temperature and maximum temperature is 0.99). (Note that the models shown in Figure 9 included a natural spline of mean temperature with 4 *df*; we observed no differences in the effects on pollutants using 2 *df* [data not shown].) The response function for temperature was rather different than the slowly increasing curve produced in our original GAM analyses (data not shown), which suggests why our air pollution effects are highly attenuated. Figure 10 shows that only one weather variable (mean temperature) greatly affected the findings for particulates and nonaccidental mortality. Mean temperature had such a profound effect on the particle associations for several possible reasons: 1) the results represent the true effect for temperature that was

missed in our original GAM analysis because concurvity among covariates was not taken into account in the backfitting algorithm; 2) temperature affects mortality but has a threshold response function (effectively zero below approximately 22°C or 25°C), so the natural spline models overfit the data and hence remove the air pollution effect; and 3) 1 and 2 are correct to varying degrees, but the observed confounding effects were accentuated because of transference of causal effects from less-precisely to more-precisely measured variables (10). Air pollution measurements were least accurate, followed by weather and time; some of the effects on air pollution could have been transferred to the other two variables.

Empirical testing of the third hypothesis is impossible, but hypothesis 2 may be correct. (We do not have extensive knowledge of the biology of acute temperature effects.) Threshold models do exist (11), but we did not have time to investigate these. Instead, we developed a series of sensitivity analyses using various functional forms for temperature assuming no effects below a certain limit. (These are not true threshold models, as the threshold is not estimated from the data.) Essentially, these models are akin to including covariates for so-called extreme temperature days, except that we modeled them as linear covariates as follows:

If $T < T_0$, $T' = 0$ and if $T \geq T_0$ then $T' = T - T_0$.

(We refer to these as *quasithreshold models*.) Two values of T_0 were used: 22°C (corresponding to a maximum temperature of between 23°C and 30°C [7% of the data above this value]) and 25°C (corresponding to a maximum temperature of between 27°C and 31°C [1.4% of the data above this

Table 4. Daily Nonaccidental Mortality: Example Showing Different Selected Temporal Filters, Calendar Year, and Weather Variables (all ages and entire time period)

Pollutant	Temporal Filter (natural spline df)	Calendar Year	Weather Variables (lag, df)
Nonaccidental Mortality			
Gaseous pollutants and COH[a]	88	Numeric	Mean temperature(0, 4) + mean relative humidity(2, 1)
Extinction	92	Numeric	Mean temperature(0, 5) + mean relative humidity(2, 1)
Predicted $PM_{2.5}$ and Sulfates from $PM_{2.5}$	74	Factor	Mean temperature(0, 5) + mean relative humidity(2, 1)
Sutton Sulfate	70	None	Mean temperature(0, 5) + mean relative humidity(2, 1)
Cancer Subgroup[b]			
Gaseous pollutants and COH[a]	68	None	Mean temperature(0, 3) + mean relative humidity(1, 2)
Predicted $PM_{2.5}$ and Sulfates from $PM_{2.5}$	23	None	Maximum temperature(0, 3) + change in pressure(0, 6)
Sutton Sulfate	50	None	Mean temperature(0, 2) + mean relative humidity(1, 2)
Acute Lower Respiratory Subgroup[b]			
Gaseous pollutants and COH[a]	43	None	Mean temperature(1, 4) + mean relative humidity(2, 0)
Predicted $PM_{2.5}$ and Sulfates from $PM_{2.5}$	52	Numeric	Dew point temperature(0, 5) + mean relative humidity(2, 0)
Sutton Sulfate	47	None	Mean temperature(1, 4) + change in pressure(0, 0)
Chronic Coronary Artery Disease Subgroup[b]			
Gaseous pollutants and COH[a]	49	Numeric	Mean temperature(0, 7) + change in pressure(0, 2)
Predicted $PM_{2.5}$ and Sulfates from $PM_{2.5}$	36	None	Mean temperature(0, 3) + mean relative humidity(2, 0)
Sutton Sulfate	42	None	Mean temperature(0, 3) + mean relative humidity(2, 0)
Congestive Heart Failure Subgroup[b]			
Gaseous pollutants and COH[a]	39	Numeric	Maximum temperature(1, 5) + mean relative humidity(1, 0)
Predicted $PM_{2.5}$ and Sulfates from $PM_{2.5}$	37	Factor	Maximum temperature(1, 7) + mean relative humidity(0, 0)
Sutton Sulfate	37	Factor	Maximum temperature(1, 4) + mean relative humidity(1, 0)
Any Coronary Artery Disease Subgroup[b]			
Gaseous pollutants and COH[a]	27	None	Maximum temperature(0, 4) + change in pressure(2, 0)
Predicted $PM_{2.5}$ and Sulfates from $PM_{2.5}$	30	None	Mean temperature(0, 3) + mean relative humidity(2, 7)
Sutton Sulfate	30	None	Mean temperature(0, 3) + mean relative humidity(2, 0)
Any Cardiovascular Disease Subgroup[b]			
Gaseous pollutants and COH[a]	58	None	Mean temperature(0, 7) + mean relative humidity(1, 0)
Predicted $PM_{2.5}$ and Sulfates from $PM_{2.5}$	52	None	Mean temperature(0, 6) + mean relative humidity(2, 0)
Sutton Sulfate	50	None	Mean temperature(0, 7) + mean relative humidity(1, 0)

[a] COH = coefficient of haze.

[b] Subgroup defined from Quebec Health Insurance Plan.
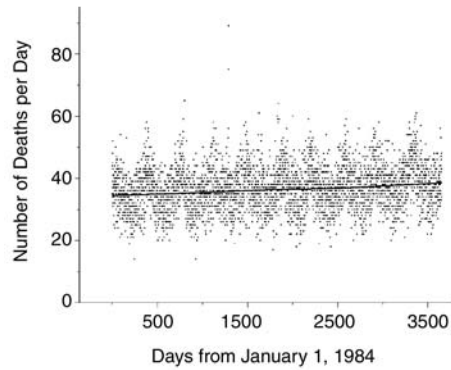
Figure 1. Time series for daily nonaccidental mortality, Montreal, 1984 to 1993.
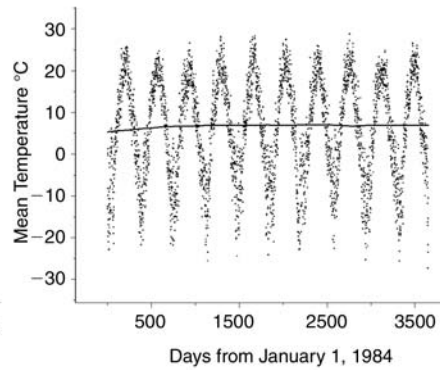
Figure 2. Time series for mean daily temperature, Montreal, 1984 to 1993.
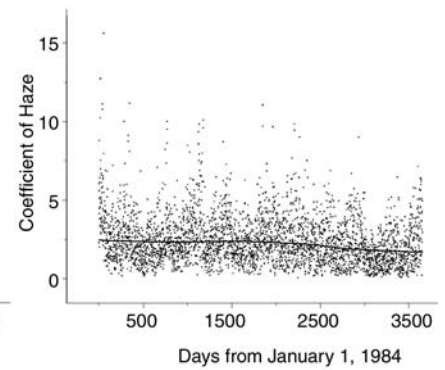
Figure 3. Time series for coefficient of haze, Montreal, 1984 to 1993.

Table 5. Daily Nonaccidental Mortality: Comparison of Original Results to Those Obtained Using 1) Stricter Convergence Criteria for GAM (original functional forms retained), 2) Natural Splines for Temporal Filter and Weather Variables, 3) Natural Splines for Temporal Filter, Weather Variables and Day of the Week. Mean Percent Change in Daily Nonaccidental Mortality for Different Measures of Particulates Evaluated at the 3-Day Mean Across Interquartile Ranges of Various Pollutants, Montreal, 1984 to 1993[a]

Cause of Death	COH[b] Original	COH[b] Reanalysis	Extinction Original	Extinction Reanalysis	Predicted PM$_{2.5}$ Original	Predicted PM$_{2.5}$ Reanalysis	Sutton Sulfate Original	Sutton Sulfate Reanalysis	Predicted Sulfate from PM$_{2.5}$ Original	Predicted Sulfate from PM$_{2.5}$ Reanalysis
1) GAM with Stricter Convergence Criteria (epsilon and bf.epsilon = 10^{-15}, maxit and bf.maxit = 1000)										
Nonaccidental deaths	1.98[c]	1.38[c]	1.67[c]	1.34[c]	2.17[c]	1.57[c]	1.29[c]	1.03[c]	1.59[c]	1.20[c]
≥65 years	2.57[c]	1.92[c]	1.96[c]	1.59[c]	2.68[c]	1.97[c]	1.77[c]	1.46[c]	2.11[c]	1.65[c]
< 65 years	0.30	−0.30	0.88	0.50	1.03	0.28	0.04	−0.36	0.27	−0.25
2) GLM with Natural Splines Using Pollutant-Specific Temporal Filters										
Nonaccidental deaths	1.98[c]	0.85	1.67[c]	0.50	2.17[c]	0.55	1.29[c]	0.27	1.59[c]	0.41
≥ 65 years	2.57[c]	1.17	1.96[c]	0.73	2.68[c]	0.84	1.77[c]	0.69	2.11[c]	0.71
< 65 years	0.30	−0.64	0.88	0.19	1.03	0.04	0.04	−0.36	0.27	−0.49
3) GLM with Natural Splines Using Pollutant-Specific Temporal Filters and Day of the week										
Nonaccidental deaths	1.98[c]	0.58	1.67[c]	0.49	2.17[c]	0.46	1.29[c]	0.29	1.59[c]	0.41
≥ 65 years	2.57[c]	0.85	1.96[c]	0.70	2.68[c]	0.70	1.77[c]	0.70	2.11[c]	0.69
< 65 years	0.30	−0.84	0.88	0.21	1.03	0.08	0.04	−0.34	0.27	−0.46

[a] The original statistical model was E[log(y_i)] = α + LOESS(i, span=x) + LOESS(year) + multiple weather variables + β(pollutant), where i is an indicator for day and x is the selected span (%). Sulfate data from the Sutton monitoring station were available only for 1986 to 1993.

[b] COH = coefficient of haze.

[c] Corrected t value >1.96.
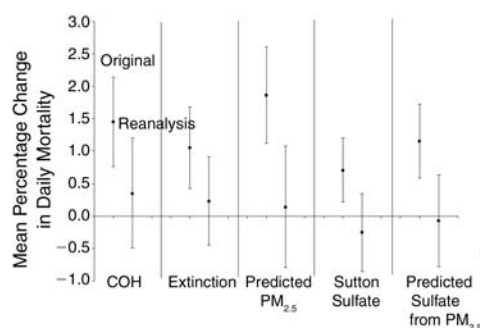
Figure 4. Daily nonaccidental deaths: Comparison of original and reanalysis (using natural splines) for all ages. Mean percent increase in daily mortality evaluated at lag 0 across interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.
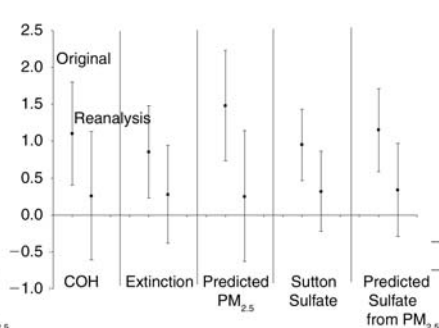
Figure 5. Daily nonaccidental deaths: Comparison of original and reanalysis (using natural splines) for all ages. Mean percent increase in daily mortality evaluated at lag 1 across interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.
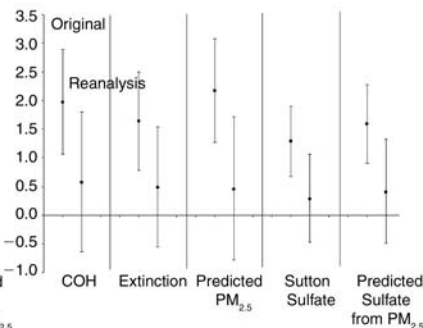
Figure 6. Daily nonaccidental deaths: Comparison of original and reanalysis (using natural splines) for all ages. Mean percent increase in daily mortality evaluated at the 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.
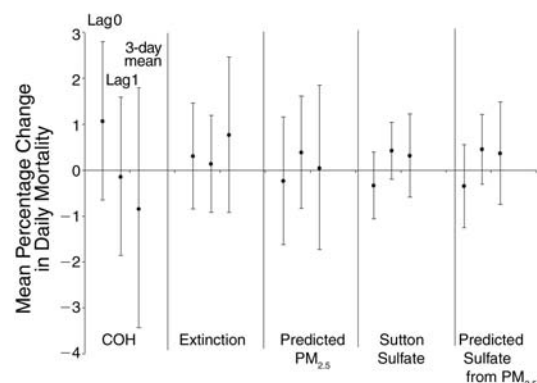
Figure 7. Daily nonaccidental deaths for the warm season: Reanalysis (using natural splines) for all ages. Mean percent increase in daily mortality evaluated at lag 0, lag 1, and the 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.
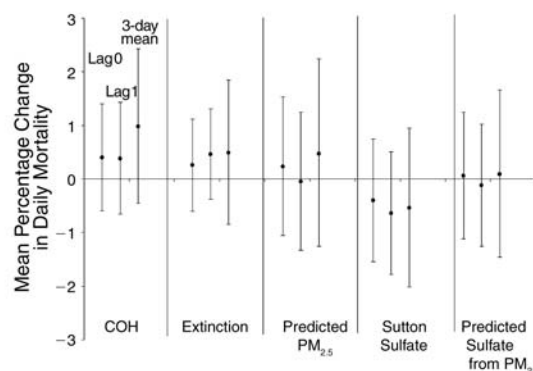
Figure 8. Daily nonaccidental deaths for the cold season: Reanalysis (using natural splines) for all ages. Mean percent increase in daily mortality evaluated at lag 0, lag 1, and the 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.

value]). Table 7 shows sensitivity analyses for COH evaluated at lag 0 days. (The AIC was adjusted to account for the implicit threshold value included in the model by adding 2 [2×1 *df*] for each implicit threshold.) The reanalysis unadjusted for weather showed a lower effect than those found in the adjusted models from the original GAM analysis (mean percent change [MPC] of 1.19% vs 1.45%). In the present reanalyses, relative humidity had little effect on air pollution and inclusion of the quasithreshold model showed only slightly attenuated effects for COH (as T_0 was reduced, so were the estimates for COH): for T_0 = 22°C, MPC = 0.89%; for T_0 = 25°C, MPC = 1.09%.

To relate our new findings to other studies, we reexpressed the MPC for a 10 µg/m^3 increase in predicted PM$_{2.5}$ (particulate matter less than 2.5 µm in diameter).

For all deaths, the percent increase in daily mortality was 0.46 per 10 µg/m^3, and for deaths among persons aged 65 years or more the percent increase in daily mortality was 0.74 per 10 µg/m^3, which is within the observed range in North America.

Subgroups Defined Using QHIP Data

For all subgroups (see part 2 of the original report [1]), we conducted analyses analogous to those for all nonaccidental mortality in which we found positive effects (cancer, acute lower respiratory diseases, chronic coronary artery disease, congestive heart failure, any coronary artery diseases, and any cardiovascular disease; see Table 22 of the original report). In addition, we conducted detailed analyses for the warm season among subjects aged 65 years or more.

Table 6. Daily Nonaccidental Mortality: Sensitivity Analyses Comparing Different Temporal Filters. Mean Percent Change in Daily Nonaccidental Mortality for Coefficient of Haze (COH) Evaluated at Lag 0 Across Interquartile Ranges, Montreal, 1984 to 1993

Model[a,b]	P value for Bartlett	AIC	Dispersion	MPC[c]	95% CI
42 *df*					
M_1	0.001	4110.05	1.107		
M_1 + COH	0.002	4103.92	1.105	1.14	0.32−1.96
M_1 + COH + ns(mt, 2 *df*)	0.008	4046.66	1.089	0.44	−0.40−1.28
M_1 + COH + ns(mt, 4 *df*)	0.040	4008.68	1.078	0.47	−0.36−1.31
88 *df*					
M_1	0.263	4044.17	1.076		
M_1 + COH	0.263	4037.22	1.073	1.19	0.37−2.02
M_1 + COH + ns(mt, 2 *df*)	0.333	3984.35	1.058	0.43	−0.42−1.27
M_1 + COH + ns(mt, 4 *df*)	0.250	3947.32	1.048	0.47	−0.37−1.32
122 *df*					
M_1	0.041	4074.34	1.074		
M_1 + COH	0.040	4066.64	1.071	1.24	0.41−2.07
M_1 + COH + ns(mt, 2 *df*)	0.039	4007.53	1.005	0.39	−0.46−1.24
M_1 + COH + ns(mt, 4 *df*)	0.028	3974.70	1.046	0.44	−0.41−1.29

[a] M_1 = Year + Day of the Week + ns(time, XX *df*).

[b] ns = natural spline, mt = mean temperature.

[c] Mean percent change across the interquartile range of COH.
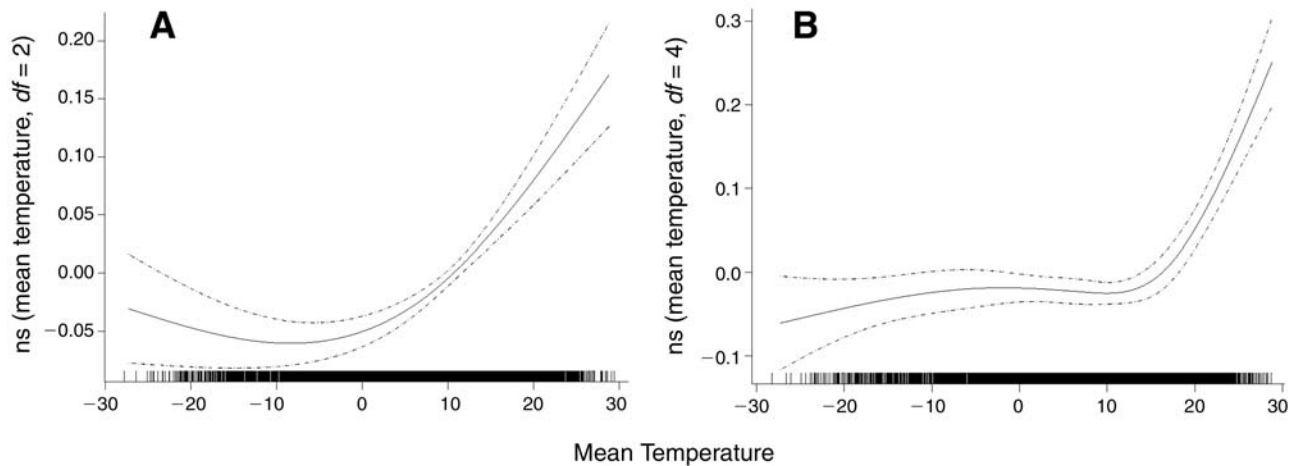
Figure 9. Daily nonaccidental mortality: Functional form for mean temperature (℃) (***A***) **using natural splines with 2 *df* and** (***B***) **using natural splines with 4 *df* (ns = natural spline).**
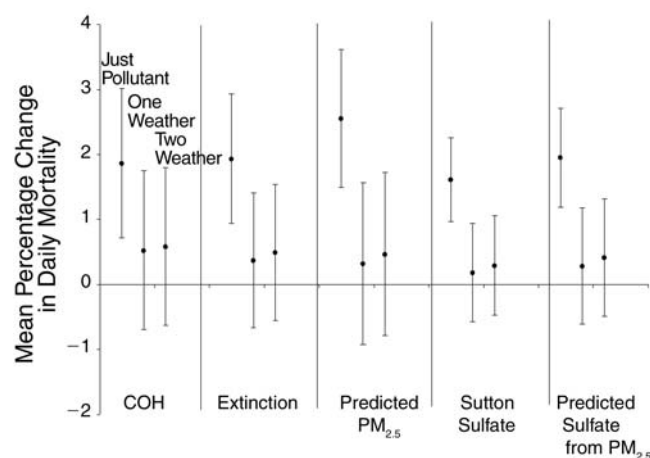
Figure 10. Daily nonaccidental deaths: Results for different weather models (using natural splines) for all ages. Mean percent increase in daily mortality evaluated at the 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.

Table 8 shows, for pollutants evaluated at the 3-day mean, that the use of the more stringent convergence criteria attenuated the effects, although in most cases the original conclusions were not altered (except for cancer and predicted $PM_{2.5}$; chronic coronary disease and all three pollutants, congestive heart failure and Sutton sulfate; any coronary disease and predicted $PM_{2.5}$; and any coronary

disease and Sutton sulfate). All analyses using natural splines showed nonsignificant effects for air pollution.

Figures 11 through 16 compare for each lag the original findings (panel A) to the results from the reanalysis (panel B) for each of these subgroups. Except for any cardiovascular disease evaluated at lag 1 (Figure 16B), all of the new results were not significantly different from no effect.

We also conducted analyses for the warm season (Figures 17–22) and found significant positive associations for congestive heart failure only (COH, extinction, predicted $PM_{2.5}$, predicted sulfate from $PM_{2.5}$; these same findings are in press in *Environmental Research*).

Tables 9 and 10 show the sensitivity analyses for the COH, evaluated at lag 0 days, for different methods of adjusting for mean temperature and relative humidity. Table 9 refers to deaths during the entire year for all age groups and Table 10 refers to deaths during the warm season among subjects aged 65 years or more. We compared models using natural splines to those assuming no temperature effects below 22°C or 25°C and no relative humidity effects below 88%. Relative humidity had little effect on estimates and larger pollutant effects were found for models using a 25°C cutoff.

Considering the cutoff model of 25°C as the most plausible, we found a significant association for the entire year (Table 9) only for any cardiovascular disease; for the warm period (Table 10), we found positive associations for chronic coronary artery disease, congestive heart failure, any coronary artery disease, and any cardiovascular disease.

Table 7. Daily Nonaccidental Mortality: Sensitivity Analyses Comparing Different Weather Models. Mean Percent Change in Daily Nonaccidental Mortality for Coefficient of Haze (COH) Evaluated at Lag 0 Across Its Interquartile Range, Montreal, 1984 to 1993

Model[a]	P value for Bartlett	AIC	Dispersion	MPC[b]	95% CI
Original analysis				1.45	0.76−2.14
Base	0.263	4037.22	1.073	1.19	0.37−2.02
Base + ns(mt, 2 *df*)	0.333	3984.35	1.058	0.43	−0.42−1.27
Base + ns(mt, 4 *df*)	0.250	3947.32	1.048	0.47	−0.37−1.32
Base + ns(RHMean, 2 *df*)	0.273	4040.61	1.074	1.10	0.25−1.95
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.336	3988.16	1.059	0.36	−0.49−1.23
Base + β=0*(mt < 22) + β*(mt − 22)	0.186	3979.33	1.057	0.89	0.07−1.71
Base + β=0*(mt < 25) + β*(mt − 25)	0.253	4016.07	1.067	1.09	0.27−1.92
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.266	4040.63	1.073	1.14	0.31−1.98
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.257	4019.39	1.067	1.04	0.21−1.88

[a] Base = COH + year + day of the week + ns(time, 88 *df*); ns = natural spline, mt = mean temperature, RHMean = mean relative humidity.

[b] Mean percent change across the interquartile range of COH.

Table 8. Subgroups Defined Using Quebec Health Insurance Plan Data: Comparison of Original Results to Those Obtained Using 1) Stricter Convergence Criteria for GAM (original functional forms retained), 2) Natural Splines for Temporal Filter, Day of the Week, and Weather Variables. Mean Percent Change (MPC) in Daily Nonaccidental Mortality for Different Measures of Particulates Evaluated at the 3-Day Mean Across Interquartile Ranges of Various Pollutants, Montreal, 1984 to 1993

Morbidity	COH[a]		Predicted PM$_{2.5}$		Sutton Sulfate	
	MPC	95% CI	MPC	95% CI	MPC	95% CI
Cancer						
Original	2.42	0.87–4.00	1.84	0.33–3.37	0.89	−0.12–1.92
Original more stringent convergence criteria	1.76	0.22–3.33	1.00	−0.50–2.52	0.52	−0.49–1.54
Reanalysis	0.46	−1.62–2.59	0.93	−0.93–2.84	0.22	−1.01–1.48
Acute lower respiratory disease						
Original	5.09	2.47–7.79	4.72	2.23–7.28	2.25	0.56–3.98
Original more stringent convergence criteria	4.18	1.58–6.84	3.67	1.21–6.20	1.68	−0.01–3.39
Reanalysis	2.20	−1.04–5.55	−0.01	−3.37–3.47	0.10	−2.05–2.29
Chronic coronary artery disease						
Original	2.62	0.53–4.75	2.20	0.14–4.31	0.63	−0.77–2.06
Original more stringent convergence criteria	1.94	−0.14–4.07	1.36	−0.70–3.45	0.22	−1.18–1.65
Reanalysis	1.10	−1.65–3.93	0.48	−2.30–3.34	−0.10	−1.81–1.64
Congestive heart failure						
Original	4.99	2.44–7.60	4.02	1.61–6.48	1.91	0.28–3.56
Original more stringent convergence criteria	4.14	1.61–6.73	2.98	0.60–5.42	1.34	−0.27–2.99
Reanalysis	2.15	−1.08–5.47	1.60	−1.62–4.92	0.65	−1.33–2.66
Any coronary artery disease						
Original	2.99	1.13–4.88	1.85	0.03–3.70	0.69	−0.55–1.95
Original more stringent convergence criteria	2.26	0.41–4.15	0.96	−0.85–2.80	0.23	−1.01–1.48
Reanalysis	1.39	−0.95–3.79	0.44	−1.97–2.91	0.08	−1.38–1.56
Any cardiovascular disease						
Original	3.65	2.23–5.09	2.76	1.40–4.15	1.16	0.23–2.09
Original more stringent convergence criteria	2.73	1.32–4.15	1.74	0.39–3.11	0.69	−0.23–1.62
Reanalysis	1.73	−0.15–3.65	1.14	−0.74–3.05	0.12	−1.04–1.30

[a] COH = coefficient of haze.
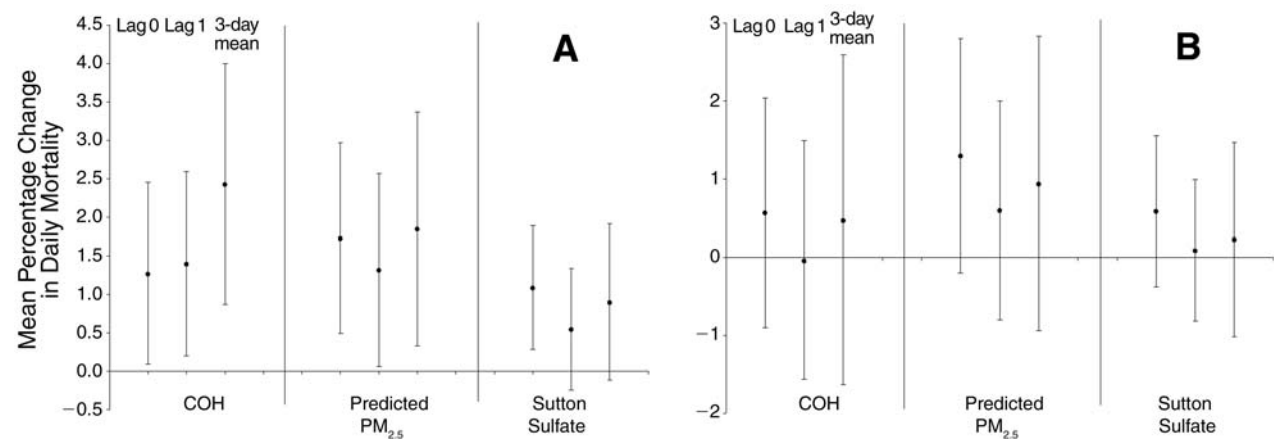
Figure 11. Cancer defined using Quebec Health Insurance Plan data: Findings for all ages, evaluated at lag 0, lag 1 and 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. *A.* Original findings. *B.* Reanalysis. COH = coefficient of haze.
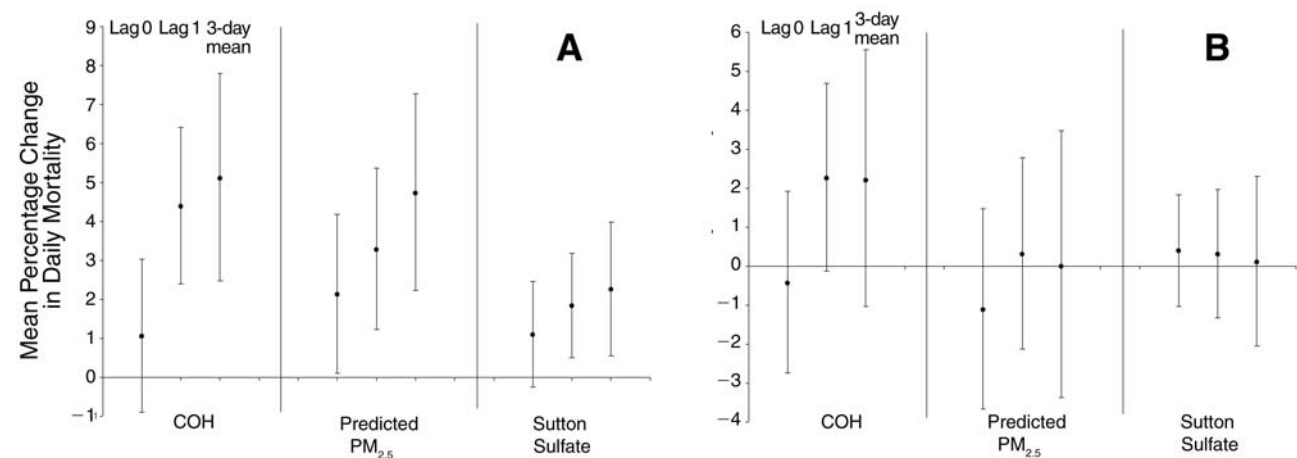
Figure 12. Acute lower respiratory disease defined using Quebec Health Insurance Plan data: Findings for all ages, evaluated at lag 0, lag 1 and 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. *A.* Original findings. *B.* Reanalysis. COH = coefficient of haze.
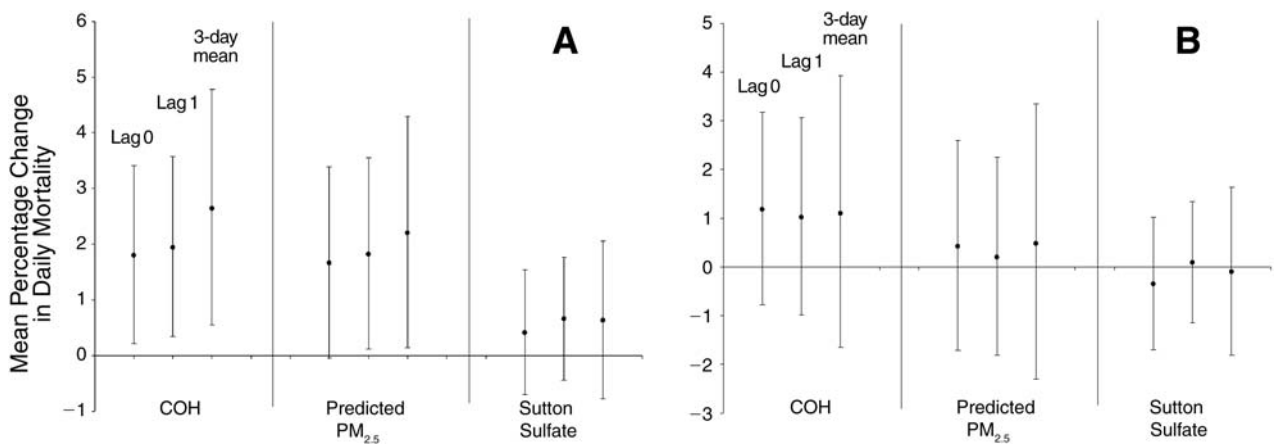
Figure 13. Chronic coronary artery disease defined using Quebec Health Insurance Plan data: Findings for all ages, evaluated at lag 0, lag 1 and 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. *A.* Original findings. *B.* Reanalysis. COH = coefficient of haze.

123
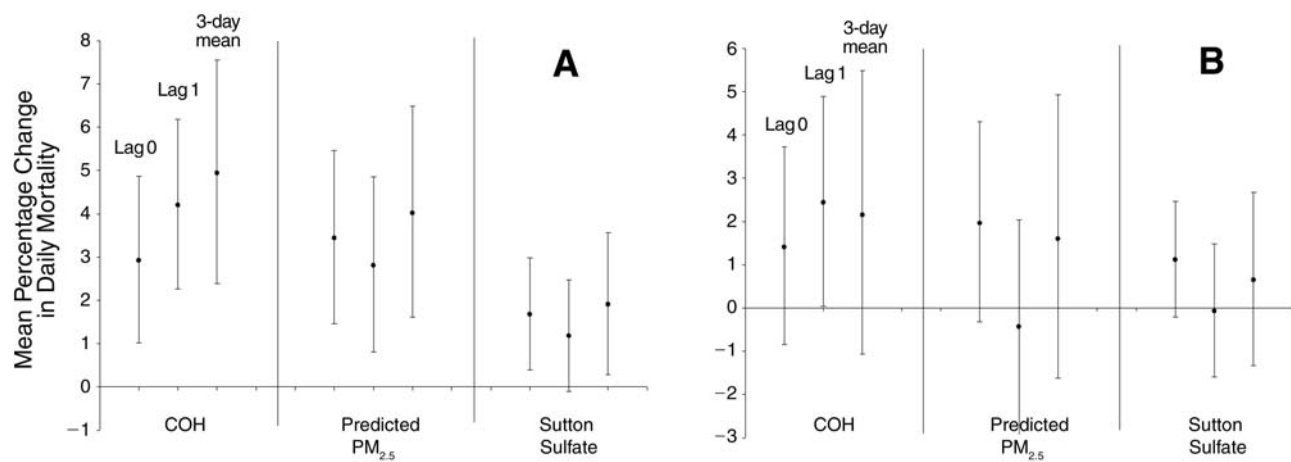
Figure 14. Congestive heart failure defined using Quebec Health Insurance Plan data: Findings for all ages, evaluated at lag 0, lag 1 and 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. *A.* Original findings. *B.* Reanalysis. COH = coefficient of haze.
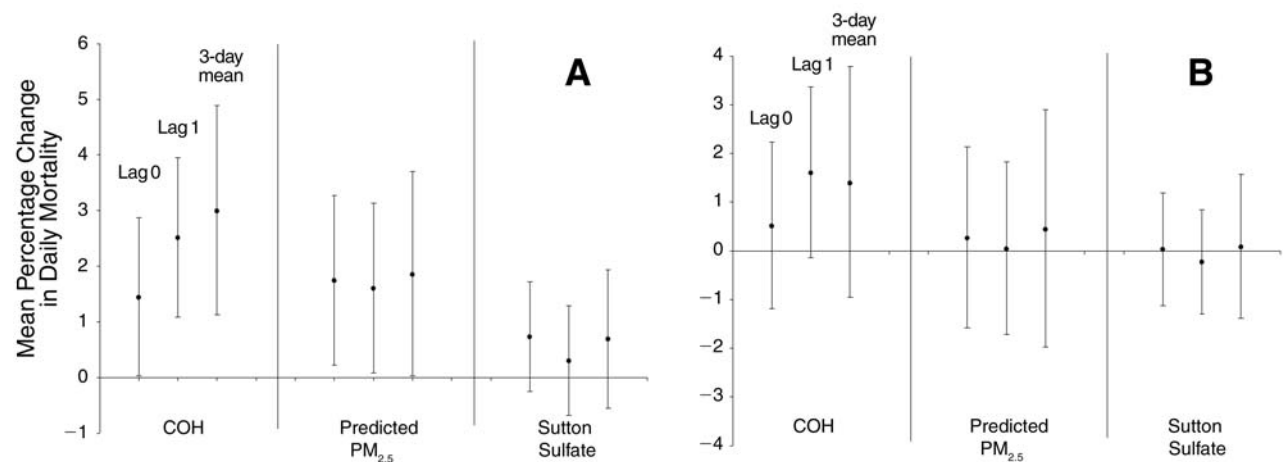
Figure 15. Any coronary artery disease defined using Quebec Health Insurance Plan data: Findings for all ages, evaluated at lag 0, lag 1 and 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. *A.* Original findings. *B.* Reanalysis. COH = coefficient of haze.
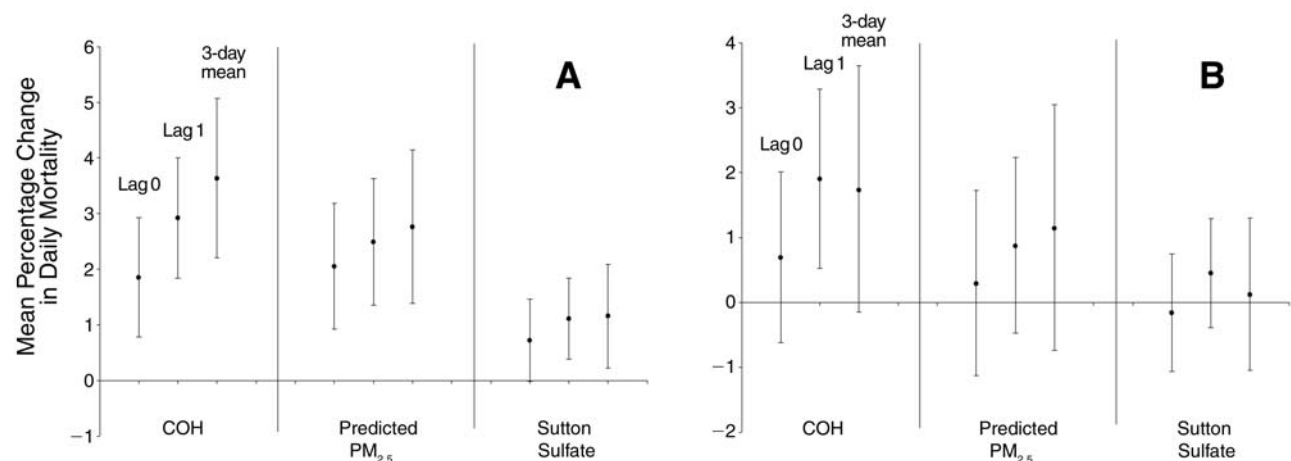
Figure 16. Any cardiovascular disease defined using Quebec Health Insurance Plan data: Findings for all ages, evaluated at lag 0, lag 1 and 3-day mean across interquartile ranges of various pollutants, Montreal, 1984 to 1993. *A.* Original findings. *B.* Reanalysis. COH = coefficient of haze.
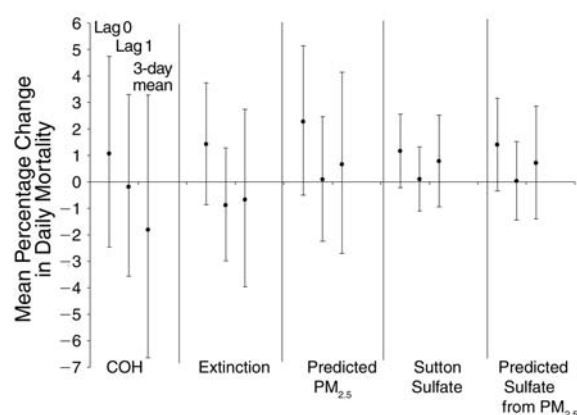
Figure 17. Cancer defined using Quebec Health Insurance Plan data: Reanalysis findings for subjects 65 years old and older, in the warm season by lag, evaluated at interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.
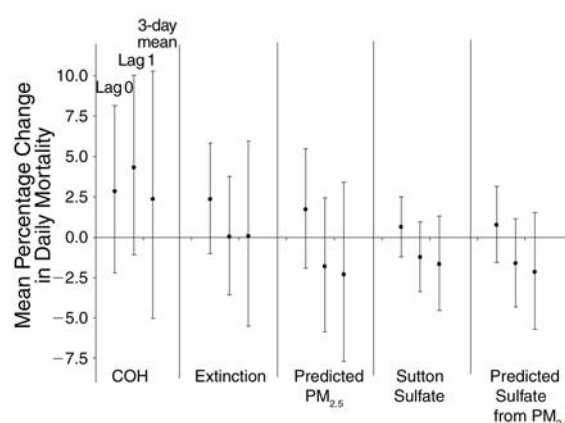
Figure 18. Acute lower respiratory disease defined using Quebec Health Insurance Plan data: Reanalysis findings for subjects 65 years old and older, in the warm season by lag, evaluated at interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.
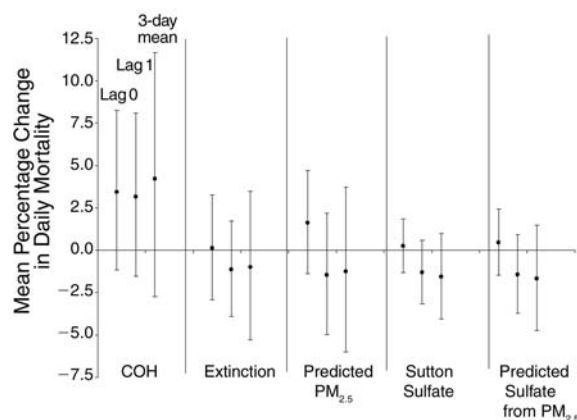

Figure 19. Chronic coronary artery disease defined using Quebec Health Insurance Plan data: Reanalysis findings for subjects 65 years old and older, in the warm season by lag, evaluated at interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.
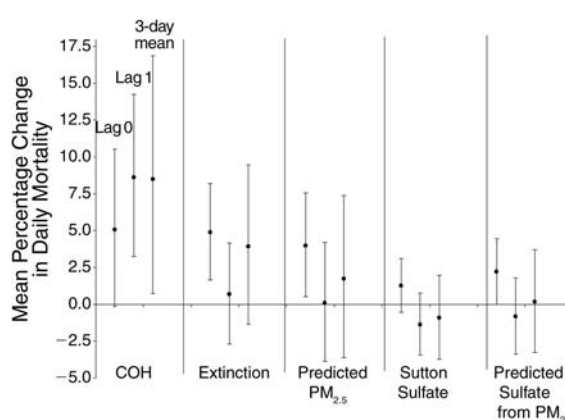
Figure 20. Congestive heart failure defined using Quebec Health Insurance Plan data: Reanalysis findings for subjects 65 years old and older, in the warm season by lag, evaluated at interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.
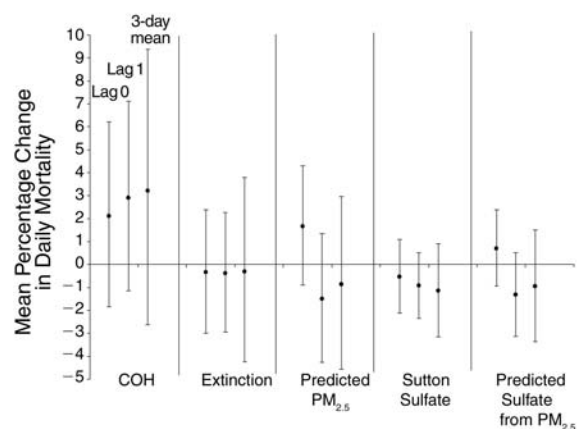

Figure 21. Any coronary artery disease defined using Quebec Health Insurance Plan data: Reanalysis findings for subjects 65 years old and older, in the warm season by lag, evaluated at interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.
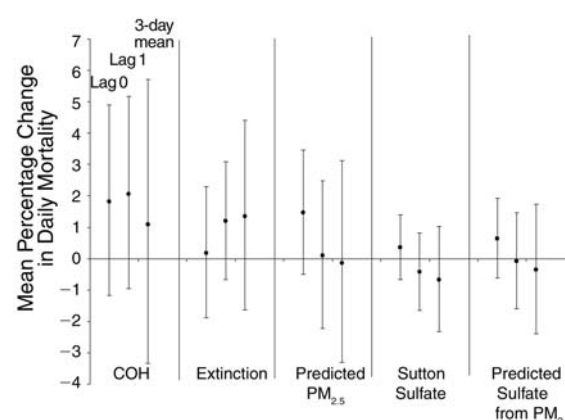
Figure 22. Any cardiovascular disease defined using Quebec Health Insurance Plan data: Reanalysis findings for subjects 65 years old and older, in the warm season by lag, evaluated at interquartile ranges of various pollutants, Montreal, 1984 to 1993. COH = coefficient of haze.

125

Table 9. Subgroups Defined Using Quebec Health Insurance Plan Data: Sensitivity Analysis for Coefficient of Haze (COH) for Entire Time Period Using Different Models for Weather. Mean Percent Change in Daily Nonaccidental Mortality Evaluated at Lag 0 Across the COH Interquartile Range, Montreal, 1984 to 1993

Model[a,b]	P value for Bartlett	AIC	Dispersion	MPC[c]	95% CI
Cancer					
Base	0.311	3853.39	1.021	1.48	0.07–2.92
Base + ns(mt, 2 *df*)	0.430	3827.68	1.012	0.53	−0.92–2.00
Base + ns(mt, 3 *df*)	0.452	3823.78	1.012	0.54	−0.91–2.02
Base + ns(mt, 4 *df*)	0.450	3825.63	1.012	0.54	−0.91–2.02
Base + ns(RHMean, 2 *df*)	0.325	3856.52	1.021	1.32	−0.13–2.79
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.437	3830.41	1.013	0.36	−1.12–1.86
Base + β=0*(mt < 22) + β*(mt − 22)	0.378	3843.13	1.017	1.22	−0.20–2.66
Base + β=0*(mt < 25) + β*(mt − 25)	0.306	3851.70	1.019	1.40	−0.01–2.84
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.313	3857.40	1.021	1.45	0.03–2.90
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.308	3855.69	1.019	1.37	−0.05–2.82
Acute Lower Respiratory					
Base	0.589	4001.31	1.014	−0.05	−2.33–2.28
Base + ns(mt, 2 *df*)	0.560	3999.10	1.013	−0.78	−3.12–1.62
Base + ns(mt, 4 *df*)	0.447	3982.47	1.009	−0.79	−3.13–1.60
Base + ns(mt(lag 1), 4 *df*)	0.524	3976.24	1.008	−0.38	−2.68–1.97
Base + ns(RHMean, 2 *df*)	0.587	4005.02	1.015	−0.13	−2.46–2.25
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.558	4002.88	1.013	−0.85	−3.23–1.60
Base + β=0*(mt < 22) + β*(mt − 22)	0.496	3990.41	1.011	−0.53	−2.81–1.80
Base + β=0*(mt < 25) + β*(mt − 25)	0.490	3988.47	1.009	−0.33	−2.61–1.99
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.584	4004.86	1.014	0.05	−2.25–2.40
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.591	4006.44	1.014	−0.10	−2.43–2.29
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.484	3992.07	1.009	−0.24	−2.53–2.11
Chronic Coronary Artery Disease					
Base	0.489	3724.28	0.965	1.51	−0.38–3.42
Base + ns(mt, 2 *df*)	0.507	3723.54	0.964	0.94	−1.00–2.92
Base + ns(mt, 4 *df*)	0.557	3713.25	0.960	0.96	−0.98–2.94
Base + ns(mt, 7 *df*)	0.537	3712.16	0.959	0.95	−0.99–2.93
Base + ns(RHMean, 2 *df*)	0.486	3728.19	0.965	1.52	−0.42–3.49
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.501	3727.31	0.964	1.02	−0.96–3.04
Base + β=0*(mt < 22) + β*(mt − 22)	0.461	3721.11	0.962	1.24	−0.65–3.17
Base + β=0*(mt < 25) + β*(mt − 25)	0.476	3723.12	0.963	1.39	−0.50–3.31
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.486	3728.17	0.965	1.54	−0.36–3.48
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.487	3730.12	0.965	1.53	−0.41–3.50
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.476	3727.02	0.963	1.42	−0.48–3.36

Table continues next page

[a] Base = COH + year + day of the week + ns(time, XX *df*).

[b] ns = natural spline, mt = mean temperature, RHMean = mean relative humidity.

[c] Mean percent change across the interquartile range of COH.

Table 9 (*Continued*). Subgroups Defined Using Quebec Health Insurance Plan Data: Sensitivity Analysis for Coefficient of Haze (COH) for Entire Time Period Using Different Models for Weather. Mean Percent Change in Daily Nonaccidental Mortality Evaluated at Lag 0 Across the COH Interquartile Range, Montreal, 1984 to 1993

Model[a,b]	P value for Bartlett	AIC	Dispersion	MPC[c]	95% CI
Congestive Heart Failure					
Base	0.989	3986.25	1.029	1.79	−0.45–4.08
Base + ns(mt, 2 *df*)	0.995	3984.41	1.028	1.03	−1.28–3.39
Base + ns(mt, 4 *df*)	0.991	3981.35	1.028	1.04	−1.27–3.41
Base + ns(mt, 5 *df*)	0.991	3983.36	1.028	1.05	−1.26–3.41
Base + ns(mt(lag 1), 5 *df*)	0.990	3980.92	1.028	1.49	−0.78–3.81
Base + ns(RHMean, 2 *df*)	0.990	3990.04	1.030	1.75	−0.55–4.09
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.995	3988.33	1.029	1.04	−1.31–3.46
Base + β=0*(mt < 22) + β*(mt − 22)	0.994	3982.40	1.027	1.42	−0.83–3.72
Base + β=0*(mt < 25) + β*(mt − 25)	0.991	3983.62	1.028	1.61	−0.63–3.90
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.989	3990.17	1.030	1.85	−0.41–4.16
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.990	3992.06	1.030	1.76	−0.54–4.11
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.992	3987.55	1.028	1.66	−0.60–3.97
Any Coronary Artery Disease					
Base	0.964	3788.03	0.999	0.83	−0.82–2.50
Base + ns(mt, 2 *df*)	0.964	3787.77	0.998	0.43	−1.25–2.14
Base + ns(mt, 4 *df*)	0.890	3779.41	0.995	0.36	−1.32–2.07
Base + ns(RHMean, 2 *df*)	0.963	3791.74	0.999	0.94	−0.76–2.66
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.963	3791.15	0.998	0.58	−1.15–2.34
Base + β=0*(mt < 22) + β*(mt − 22)	0.875	3783.35	0.996	0.55	−1.10–2.23
Base + β=0*(mt < 25) + β*(mt − 25)	0.897	3785.00	0.997	0.69	−0.95–2.37
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.964	3791.85	0.999	0.88	−0.78–2.57
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.964	3793.72	0.999	0.94	−0.75–2.67
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.897	3788.85	0.997	0.74	−0.92–2.43
Any Cardiovascular Disease					
Base	0.535	3822.63	1.015	1.39	0.12–2.67
Base + ns(mt, 2 *df*)	0.449	3809.14	1.011	0.70	−0.60–2.03
Base + ns(mt, 4 *df*)	0.312	3796.59	1.007	0.70	−0.60–2.02
Base + ns(mt, 7 *df*)	0.341	3788.37	1.005	0.68	−0.63–1.99
Base + ns(RHMean, 2 *df*)	0.520	3825.05	1.015	1.34	0.03–2.66
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.433	3811.73	1.011	0.69	−0.65–2.04
Base + β=0*(mt < 22) + β*(mt − 22)	0.260	3811.07	1.011	1.13	−0.14–2.42
Base + β=0*(mt < 25) + β*(mt − 25)	0.300	3814.76	1.011	1.27	0.00–2.56
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.515	3825.99	1.015	1.46	0.18–2.76
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.521	3827.14	1.015	1.34	0.04–2.67
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.288	3818.18	1.011	1.34	0.06–2.64

[a] Base = COH + year + day of the week + ns(time, XX *df*).

[b] ns = natural spline, mt = mean temperature, RHMean = mean relative humidity.

[c] Mean percent change across the interquartile range of COH.

Table 10. Subgroups Defined Using Quebec Health Insurance Plan Data: Sensitivity Analysis for Coefficient of Haze (COH) for Warm Season, Subjects Aged 65 Years or More, Using Different Models for Weather. Mean Percent Change in Daily Nonaccidental Mortality Evaluated at Lag 0 Across the COH Interquartile Range, Montreal, 1984 to 1993

Model[a,b]	P value for Bartlett	AIC	Dispersion	MPC[c]	95% CI
Cancer					
Base	0.755	1945.82	1.011	2.77	−0.44–6.07
Base + ns(mt, 2 *df*)	0.772	1933.88	1.004	0.85	−2.58–4.39
Base + ns(mt, 4 *df*)	0.779	1937.39	1.005	0.83	−2.60–4.38
Base + ns(RHMean, 2 *df*)	0.754	1949.59	1.012	2.77	−0.47–6.12
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.771	1937.87	1.006	0.83	−2.63–4.41
Base + β=0*(mt < 22) + β*(mt − 22)	0.759	1941.35	1.007	1.90	−1.33–5.24
Base + β=0*(mt < 25) + β*(mt − 25)	0.755	1945.60	1.009	2.47	−0.74–5.78
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.758	1947.50	1.010	2.93	−0.28–6.25
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.756	1949.25	1.011	2.80	−0.44–6.15
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.759	1947.37	1.008	2.63	−0.59–5.95
Acute Lower Respiratory					
Base	0.497	2124.84	1.033	3.90	−1.20–9.26
Base + ns(mt, 2 *df*)	0.375	2106.69	1.022	1.76	−3.58–7.39
Base + ns(mt, 4 *df*)	0.360	2109.71	1.023	1.77	−3.58–7.42
Base + ns(mt(lag 1), 4 *df*)	0.395	2102.04	1.020	2.92	−2.16–8.28
Base + ns(RHMean, 2 *df*)	0.508	2127.33	1.034	3.47	−1.66–8.86
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.391	2109.47	1.024	1.33	−4.04–7.01
Base + β=0*(mt < 22) + β*(mt − 22)	0.410	2117.56	1.029	2.31	−2.80–7.70
Base + β=0*(mt < 25) + β*(mt − 25)	0.385	2113.25	1.026	2.90	−2.17–8.24
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.496	2128.86	1.034	3.94	−1.17–9.31
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.515	2129.07	1.034	3.49	−1.64–8.88
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.384	2117.29	1.026	2.94	−2.15–8.29
Chronic Coronary Artery Disease					
Base	0.322	2078.21	1.036	6.83	2.45–11.40
Base + ns(mt, 2 *df*)	0.237	2061.62	1.025	4.04	−0.59–8.88
Base + ns(mt, 4 *df*)	0.255	2060.19	1.023	3.67	−0.95–8.50
Base + ns(RHMean, 2 *df*)	0.314	2081.17	1.036	7.15	2.71–11.79
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.235	2065.21	1.026	4.31	−0.38–9.22
Base + β=0*(mt < 22) + β*(mt − 22)	0.260	2075.62	1.032	5.82	1.40–10.42
Base + β=0*(mt < 25) + β*(mt − 25)	0.295	2079.26	1.034	6.49	2.10–11.06
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.320	2082.16	1.036	6.90	2.51–11.48
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.314	2083.87	1.037	7.11	2.67–11.75
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.295	2083.22	1.035	6.55	2.15–11.14

Table continues next page

[a] Base = COH + year + day of the week + ns(time, XX *df*).

[b] ns = natural spline, mt = mean temperature, RHMean = mean relative humidity.

[c] Mean percent change across the interquartile range of COH.

Table 10 (*Continued*). Subgroups Defined Using Quebec Health Insurance Plan Data: Sensitivity Analysis for Coefficient of Haze (COH) for Warm Season, Subjects Aged 65 Years or More, Using Different Models for Weather. Mean Percent Change in Daily Nonaccidental Mortality Evaluated at Lag 0 Across the COH Interquartile Range, Montreal, 1984 to 1993

Model[a,b]	P value for Bartlett	AIC	Dispersion	MPC[c]	95% CI
Congestive Heart Failure					
Base	0.775	2153.78	1.064	8.11	3.21–13.25
Base + ns(mt, 2 *df*)	0.791	2135.34	1.056	5.06	−0.13–10.51
Base + ns(mt, 4 *df*)	0.798	2138.50	1.057	4.85	−0.34–10.31
Base + ns(RHMean, 2 *df*)	0.782	2156.99	1.065	8.27	3.30–13.48
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.796	2139.18	1.057	5.20	−0.05–10.73
Base + β=0*(mt < 22) + β*(mt − 22)	0.760	2150.14	1.061	6.90	1.97–12.07
Base + β=0*(mt < 25) + β*(mt − 25)	0.761	2152.99	1.062	7.63	2.73–12.77
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.774	2157.58	1.064	8.20	3.29–13.35
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.773	2159.77	1.065	8.23	3.26–13.44
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.760	2156.82	1.062	7.72	2.80–12.88
Any Coronary Artery Disease					
Base	0.465	2074.41	1.061	5.27	1.52–9.16
Base + ns(mt, 2 *df*)	0.415	2051.83	1.048	2.51	−1.45–6.63
Base + ns(mt, 4 *df*)	0.445	2052.47	1.048	2.29	−1.67–6.41
Base + ns(RHMean, 2 *df*)	0.456	2075.76	1.061	5.72	1.91–9.67
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.395	2053.94	1.048	2.95	−1.07–7.13
Base + β=0*(mt < 22) + β*(mt − 22)	0.430	2069.35	1.056	4.26	0.49–8.18
Base + β=0*(mt < 25) + β*(mt − 25)	0.446	2074.17	1.059	4.93	1.18–8.82
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.468	2076.77	1.061	5.45	1.68–9.35
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.455	2077.82	1.061	5.72	1.91–9.67
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.450	2076.58	1.059	5.11	1.34–9.01
Any Cardiovascular Disease					
Base	0.418	2054.88	1.073	4.67	1.83–7.59
Base + ns(mt, 2 *df*)	0.356	2017.43	1.051	1.93	−1.07–5.02
Base + ns(mt, 4 *df*)	0.382	2015.47	1.050	1.67	−1.33–4.75
Base + ns(RHMean, 2 *df*)	0.416	2056.00	1.072	4.78	1.90–7.74
Base + ns(mt, 2 *df*) + ns(RHMean, 2 *df*)	0.340	2020.00	1.052	2.03	−1.00–5.15
Base + β=0*(mt < 22) + β*(mt − 22)	0.346	2035.77	1.062	3.55	0.69–6.48
Base + β=0*(mt < 25) + β*(mt − 25)	0.388	2048.46	1.067	4.25	1.41–7.17
Base + β=0*(RHMean < 88) + β*(RHMean − 88)	0.422	2056.06	1.072	4.83	1.98–7.76
Base + RHMean + β=0*(RHMean < 88) + β*(RHMean − 88)	0.425	2058.17	1.072	4.77	1.89–7.73
Base + β=0*(mt < 25) + β*(mt − 25) + β=0*(RHMean < 88) + β*(RHMean − 88)	0.368	2049.73	1.066	4.41	1.56–7.34

[a] Base = COH + year + day of the week + ns(time, XX *df*).

[b] ns = natural spline, mt = mean temperature, RHMean = mean relative humidity.

[c] MPC = Mean percent change across the interquartile range of COH.

DISCUSSION

This reanalysis showed clearly that the GAMs have not performed adequately, in that they did not account correctly for strong correlations among nonlinear covariates. The use of the tighter convergence criteria in this data set and in others with similar levels of correlations will not resolve these problems. Thus, the findings from this reanalysis with regard to tighter convergence criteria for the GAMs should not be trusted.

The main issue of this reanalysis relates to how one should model weather. In the original analysis, we established a goodness-of-fit criterion for selecting weather variables that, in this reanalysis, we applied routinely as our primary analysis. Previously, we were not overly concerned with the functional form for the weather terms, as weather did not affect the finding appreciably. This lack of concern now appears incorrect, as the parametric models that we are using now are correctly handling the complex correlations in our data.

With the mostly nonsignificant findings (except for congestive heart failure), arising not from filtering the data but from the apparent confounding by temperature, it is critically important to ensure that these results are correct. It is thus worthwhile to reconsider the methodology employed for these adjustments. In particular, in multivariate modeling, reliance on *P* values or other goodness-of-fit tests is not an appropriate method for selecting risk factors as potential confounding variables. Rather, one should attempt to use prior knowledge of risk factors to construct the most accurate statistical model possible. Temperature appears to be a risk factor for daily mortality, especially in fragile populations, but simply fitting a natural spline to the data may be incorrect if the effects only occur in higher temperature ranges. The data from Montreal showed that most confounding occurs in the midrange of temperature (~15–22°C). Models that account for multiple unusually warm days might better reflect actual physiological mechanisms. (In Montreal, because of centralized heating in most homes and a lack of air conditioning, the warm season is of greatest concern.) In this reanalysis, there was insufficient time to develop such models, but our sensitivity analyses using the quasithreshold models showed that air pollution effects increased when the temperature threshold was increased.

In conclusion, although our primary results were based on fitting natural splines to weather variables, which was the basis of our a priori modeling strategy, these results should be considered final. The most credible findings may lie somewhere between the results of our quasithreshold models and the use of parametric smoothers.

REFERENCES*

1. **Goldberg MS, Bailar JC III, Burnett RT, Brook JR, Tamblyn R, Bonvalot Y, Ernst P, Flegel KM, Singh RK, and Valois M-F. 2000. Identifying Subgroups of the General Population That May Be Susceptible to Short-Term Increases in Particulate Air Pollution: A Time-Series Study in Montreal, Quebec. Research Report 97. Health Effects Institute, Cambridge MA.**

2. **Goldberg MS, Burnett RT, Bailar JC III, Brook J, Bonvalot Y, Tamblyn R, Singh R, and Valois M-F. 2001. The association between daily mortality and ambient air particle pollution in Montreal, Quebec: 1. Nonaccidental mortality. Environ Res 86:12–25.**

3. **Goldberg MS, Burnett RT, Bailar JC III, Brook J, Bonvalot Y, Tamblyn R, Singh R, Valois M-F, Vincent R. 2001. The association between daily mortality and ambient air particle pollution in Montreal, Quebec: 2. Cause-specific mortality. Environ Res 86:26–36.**

4. **Goldberg MS, Burnett RT, Bailar JC III, Tamblyn R, Ernst P, Flegel K, Brook J, Bonvalot Y, Singh R, Valois M-F, Vincent R. 2001. Identification of persons with cardiorespiratory conditions who are at risk of dying from the acute effects of ambient air particles. Environ Health Perspect 109(Suppl 4):487–494.**

5. **Goldberg MS, Burnett RT, Brook J, Bailar JC III, Valois M-F, Vincent R. 2001. Associations between daily cause-specific mortality and concentrations of ground-level ozone in Montreal, Quebec. Am J Epidemiol 154:817–826.**

6. Priestly, M. B. Spectral Analysis and Time Series. Academic Press, London:1981

7. Dominici, F., McDermott, A., Zeger, S. L., and Samet, J. M. On generalized additive models in time series studies of air pollution and health. Am J Epidemiol (156):1-11,2002

8. Ramsay, T. O., Burnett, R. T., and Krewski, D. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology (14):18-23,2003

* Bold type identifies publications containing the original analyses revised in this short communication report.

9. McCullagh, P. and Nelder, J. A. Generalized Linear Models. (2nd ed). Chapman and Hall, London:1989

10. Zidek, J. V., Wong, H., Le, N. D., and Burnett, R. Causality, measurement error and multicollinearity in epidemiology. Environmetrics (7):441–451,1996

11. Cakmak, S., Burnett, R. T., and Krewski, D. Methods for detecting and estimating population threshold concentrations for air pollution-related mortality with exposure measurement error. Risk Anal (19):487–496,1999

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
COH	coefficient of haze
df	degrees of freedom
EPA	Environmental Protection Agency (US)
GAM	generalized additive model
GLM	generalized linear model
LOESS	locally weighted smoothers
MPC	mean percent change
ns	natural spline
$PM_{2.5}$	particulate matter less than 2.5 μm in diameter
QHIP	Quebec Health Insurance Plan

APPENDIX A.

Table A.1. Distribution of Mean Daily Environmental Pollutants Averaged over All Monitoring Stations, Montreal, from 1984 to 1993[a]

	Units	Monit-oring Stations (*n*)	Days of Measure-ments (*n*)	Mean	SD	Mini-mum	Percentiles 25th	50th	75th	100th	IQR[b]
TSP	μg/m³	19	603	53.1	22.6	14.6	37.0	48.7	65.6	211.1	28.57
PM_{10}	μg/m³	2	624	32.2	17.6	6.5	19.7	28.5	41.1	120.5	21.32
$PM_{2.5}$	μg/m³	2	636	17.4	11.4	2.2	9.4	14.7	21.9	72.0	12.51
Sulfate from PM_{10}	μg/m³	2	437	4.7	4.4	0.3	1.9	3.6	5.7	30.7	3.84
Sulfate from $PM_{2.5}$	μg/m³	2	446	4.3	4.2	0.2	1.6	3.1	5.1	29.2	3.51
Sulfate from TSP	μg/m³	13	607	4.3	2.9	0.3	2.3	3.6	5.3	19.2	3.02
Sutton Sulfate[c]	μg/m³	1	2680	3.3	3.6	0	1.3	2.2	3.8	30.0	2.50
COH	0.1 COH units/327.8 linear m	11	3653	2.4	1.5	0.1	1.3	2.1	3.2	15.6	1.85
Extinction		1	3454	0.15	0.10	0.01	0.06	0.15	0.17	1.87	0.11
SO_2	μg/m³	13	3653	17.8	11.2	3.9	10.3	14.6	21.8	105.7	11.50
NO_2	μg/m³	8	3653	41.7	15.4	8.8	30.9	39.5	50.2	143.5	19.34
NO	μg/m³	8	3653	41.8	29.0	2.7	21.9	34.8	52.3	281.4	30.41
CO	ppm	12	3653	0.8	0.5	0.1	0.5	0.7	1.0	5.1	0.5
O_3	μg/m³	9	3653	29.0	17.1	2.8	16.6	26.0	37.9	163.9	21.34

[a] TSP = total suspended particles; PM_{10} = particulate matter < 10 μm in diameter; COH = coefficient of haze; SO_2 = sulfur dioxide; NO_2 = nitrogen dioxide; NO = nitrogen oxide; CO = carbon monoxide; O_3 = ozone.

[a] IQR = interquartile range.

[c] From 1986 to 1993.

Daily Mortality and Air Pollution in The Netherlands

Gerard Hoek

ABSTRACT

We recently reported an association between daily fluctuations in ambient air pollution and daily mortality in the Netherlands. To adjust for long-term time trends and seasonal variation, we used generalized additive models (GAMs*). The use of GAM has been questioned because of inappropriate convergence criteria and underestimation of standard errors. We performed a reanalysis of the data to test whether the conclusions were sensitive to the convergence criteria in S-Plus and replacement of the GAM model by a natural spline (ns) function, which is assumed to give correct standard errors.

Air pollution effect estimates were in general very similar between the two GAMs (default and strict convergence criteria) and natural spline models. Standard errors of the natural spline model were between 10% and 25% higher. The specification of stricter convergence criteria and the use of natural splines instead of GAM did not change any of the conclusions drawn in the two original papers.

INTRODUCTION

We recently published two papers showing an association between daily fluctuations in ambient air pollution and daily mortality in the Netherlands (Hoek et al 2000, 2001). To adjust for long-term time trends and seasonal variation, we used GAM. The use of GAM has been questioned because of inappropriate convergence criteria and underestimation of standard errors (Dominici et al 2002; Ramsay et al 2003). In the National Morbidity, Mortality, and Air Pollution Study (NMMAPS); the estimated air pollution effect decreased with more appropriate convergence

criteria and using a parametric model to adjust for nonlinear relations (Dominici et al 2002).

Key findings of the reported studies were that both particulate matter and gaseous air pollutants were associated with daily mortality. The effect of ozone was independent of that of other air pollutants. Indicators for fine particles were more strongly associated with mortality than particulate matter of 10 μm mass median aerodynamic diameter (PM_{10}). Higher effect estimates were found for the summer season than for the winter season for all pollutants (Hoek et al 2000). Larger relative risks (RRs) were found for specific cardiovascular causes of death such as arrhythmia, heart failure and thrombosis (Hoek et al 2001) than for all cardiovascular causes of death combined.

We performed a reanalysis of these data to test whether the conclusions were sensitive to use of the default convergence criteria in S-Plus and replacement of GAM by a parametric model (natural spline).

METHODS

The data and the confounder model development has been described in detail (Hoek et al 2000). Briefly, we studied daily (cause-specific) mortality from 1986 to 1994 for regions of the Netherlands and the entire Netherlands combined. This paper focuses on entire Netherlands analyses only. Daily air pollution data were available for gaseous air pollutants (ozone [O_3], sulfur dioxide [SO_2], nitrogen dioxide [NO_2], carbon monoxide [CO]) and black smoke (BS) for the entire study period. Daily data for PM_{10} and fine particle components sulfate (SO_4^{2-}) and nitrate (NO_3^-) were available from 1992 to 1994. Daily data on temperature and relative humidity were available for the entire study period.

To adjust for nonlinear relation with confounders, generalized additive models were used. The original confounder model was developed based upon a priori selected plausible spans for the locally weighted smoothers (LOESS) function (eg, 0.01 to 0.10 for date). We used the Akaike

---

information criterion (AIC) to select the span but checked plots of the partial autocorrelation of the residuals and plots of predicted values versus predictor for potential overadjustment. The final confounder model consisted of:

- LOESS function of date with a span of 0.02 for total mortality and 0.03 for respiratory and cardiovascular mortality (94 and 60 degrees of freedom for nine years, thus a fairly tight control for season was selected); for components with fewer data, the span was adjusted accordingly (eg, 0.06 for PM_{10} and total mortality);

- three linear influenza variables (average lag 0–6 days, lag 7–13 days, and lag 14–20 days);

- two bilinear temperature (T) variables: WARM = 0 below 14°C and T − 14 above 14°C and COLD = 0 above 14°C and 14 − T below 14°C; WARM of the same day and COLD lagged three days were included as these lags showed the strongest association with mortality;

- B-spline of relative humidity with one knot at 80% (bilinear function); and

- indicator variables for day of the week and holidays.

All original calculations were performed using GAM implemented in S-Plus 4.0 for Windows. The default convergence criterion (0.001) and number of iterations (10) were used. Results from this procedure is referred to as *GAM default*. For the sensitivity analysis, the convergence criterion for the backfitting and the local scoring procedures was set to 0.00000001 and the number of iterations to 1000 for both procedures. Results from this procedure are referred to as *GAM strict*. As an additional sensitivity analysis, we replaced the LOESS function of date by a natural spline of date with the same degree of freedom as the LOESS function. Strict convergence criteria were used. The degree of freedom from the LOESS model was determined by the summary function in S-Plus. Results from this procedure are termed *Spline*. All sensitivity analyses were conducted with S-Plus 2000 for Windows.

RESULTS

Using default criteria, the procedure stopped after 7 to 10 iterations. Differences of the deviance of 1 to 3 units were quite typical, suggesting too early termination. Using more strict criteria, the procedure stopped after 15 to 30 iterations. Typically, the deviance after the first loop was already within 0.5 units from the deviance after the final loop, a result from the large number of iterations in the inner loop. The natural spline model converged in three to four loops. We noticed that with missing values,

S-Plus did not use all the degrees of freedom supplied to the ns(date) function, probably because knots were placed at missing values. This was a problem for the PM_{10} analyses where we had complete daily data for the last three years of the nine-year study period. The problem was solved by using a separate dataset of the last three years.

The deviance of GAMs with stricter convergence criteria was substantially lower than the deviance from models using default criteria (Table 1). Deviance of the natural spline model was higher than that of GAM, in agreement with the more flexible nature of the LOESS function. Plots of the LOESS and natural spline function of date showed similar patterns (Figure 1), with slightly larger amplitude for the spline function. The plot documents that with a

Table 1. Deviances of Different Models for Total Mortality, The Netherlands 1986–1994

	Lag	GAM default	GAM strict	Natural Spline
PM_{10} [a]	0-6	1159.3	1153.8	1171.5
O_3	1	3505.8	3483.3	3506.2
NO_2	0-6	3494.3	3476.1	3506.8

[a] Daily data from 1992–1994 only.
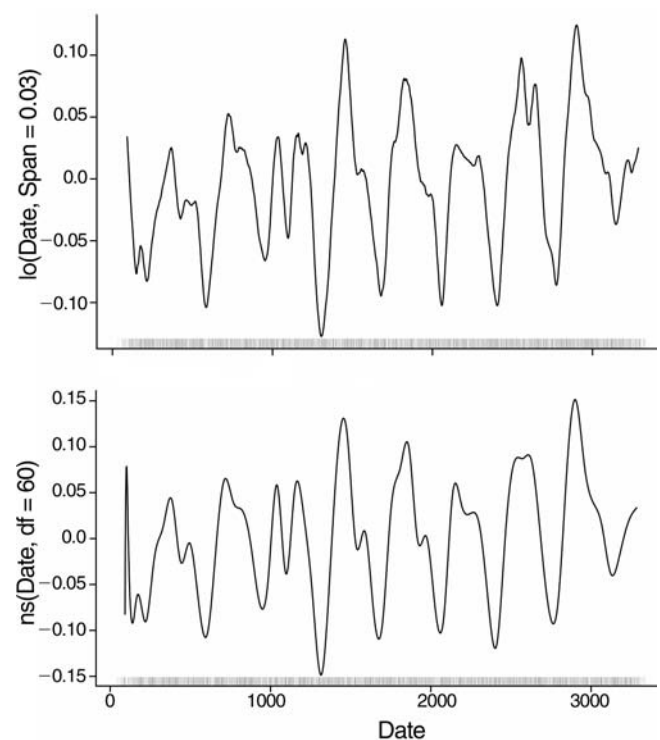
Figure 1. Predicted seasonal variation of cardiovascular mortality by LOESS (lo) and natural spline (ns) models (*df* = degree of freedom).

fairly strict control for seasonality was selected. Coefficients from temperature and influenza did not differ substantially between the LOESS and spline model.

Effect estimates for all-cause mortality were very similar in the three different models (Table 2). For most pollutants, effect estimates were slightly lower in the natural spline model compared to GAMs. For ozone, effect estimates were higher in the spline model and GAM strict, especially for the weekly average. Consistently across models, standard errors of GAMs (both default and strict) were 10% to 25% lower than those of the corresponding natural spline model. For two cases with some decrease in effect estimate in the spline model (weekly average of SO_2 and NO_2), we checked whether increasing the degree of freedom for the spline model (such that approximately the same deviance as that of the LOESS model was achieved) affected the results. Regression slopes (SE) for NO_2 were 0.000877 (0.000130) for GAM strict, 0.000648 (0.000172) for the spline model with the same degree of freedom, and 0.000737 (0.000178) for the spline model with approximately the same deviance. Corresponding numbers for SO_2 were 0.000820 (0.000130), 0.000614 (0.000156), and 0.000644 (0.000157). Thus, only a small fraction of the difference between the GAM and the spline model was due to the tighter fit of GAM.

Effect estimates for cardiovascular, chronic obstructive pulmonary disease, and pneumonia mortality were similar in the three different models (Table 3). Differences between models were even smaller than for all-cause mortality, possibly related to the larger span selected for cause-specific analyses. As in the original paper, the largest effect estimates were found for pneumonia deaths. Effect estimates for specific cardiovascular causes of death were similar across the three models (Table 4). As in the original analyses, RR estimates were substantially higher for heart failure, arrhythmia and thrombosis compared to all cardiovascular disease combined. Note that the RR estimates for cardiovascular disease combined in the original paper (Hoek et al 2001) are slightly different from the ones presented here and reported in (Hoek et al 2000). The original RRs for cardiovascular disease combined are from a report to the Ministry (Hoek et al 1997). For this report, adjustment for influenza was performed differently: only influenza of the current day was taken into account, which turned out to be insufficient. Little change occurred in the effect estimates from two pollutant models (Table 5). No conclusion about independent effects was altered by model choice.

Finally I checked whether the small changes that were found were due to the use of only one LOESS function in the

Table 2. Sensitivity of Total Mortality Associations with Air Pollution to Confounder Model Specification

Pollutant	GAM Default		GAM Strict		Natural Spline		N
	RR	95% CI	RR	95% CI	RR	95% CI	
PM_{10} lag1	1.018	1.003,1.034	1.019	1.003,1.034	1.018	1.002,1.035	1092
PM_{10} avg	1.023	1.004,1.041	1.023	1.005,1.041	1.019	0.998,1.040	1081
O_3 lag1	1.034	1.020,1.049	1.040	1.025,1.054	1.043	1.024,1.062	3196
O_3 avg	1.017	1.002,1.032	1.042	1.026,1.057	1.059	1.031,1.087	3191
BS lag1	1.020	1.010,1.030	1.019	1.010,1.029	1.019	1.009,1.030	3267
BS avg	1.028	1.017,1.038	1.026	1.015,1.036	1.022	1.009,1.035	3267
SO_2 lag1	1.027	1.017,1.037	1.026	1.016,1.036	1.023	1.012,1.034	3196
SO_2 avg	1.037	1.026,1.048	1.033	1.022,1.045	1.025	1.012,1.038	3191
NO_2 lag1	1.029	1.020,1.038	1.028	1.019,1.037	1.025	1.015,1.035	3184
NO_2 avg	1.031	1.023,1.040	1.027	1.018,1.035	1.020	1.009,1.031	3177
CO lag1	1.035	1.018,1.052	1.033	1.016,1.050	1.031	1.012,1.050	2158
CO avg	1.046	1.025,1.068	1.041	1.020,1.062	1.042	1.017,1.068	2153
SO_4^{2-} lag1	1.032	1.006,1.059	1.033	1.007,1.062	1.030	1.002,1.060	876
SO_4^{2-} avg	1.019	0.995,1.043	1.022	0.998,1.046	1.016	0.989,1.045	871
NO_3^- lag1	1.041	1.014,1.069	1.042	1.015,1.070	1.041	1.011,1.072	876
NO_3^- avg	1.029	1.005,1.053	1.030	1.006,1.054	1.024	0.996,1.053	871
Sec lag1	1.043	1.014,1.073	1.044	1.015,1.074	1.042	1.010,1.075	876
Sec avg	1.026	1.001,1.052	1.028	1.003,1.054	1.022	0.993,1.052	871

Note: The ranges for RR calculation were (weekly average in micrograms per cubic meter in parentheses) 100 (80) for PM_{10}; 50 (40) for BS and SO_2; 50 (30) for NO_2; 150 (120) for O_3; 1500 (1200) for CO; 25 (15) for sulfate and nitrate; 50 (30) for the sum of sulfate and nitrate. Sec = sum of sulfate and nitrate as an approximation of secondary aerosol; Avg = average of lag 0–6 days.

Table 3. Sensitivity of Cause-Specific Mortality Analysis with Air Pollution to Confounder Model Specification

Pollutant[a]	GAM Default		GAM Strict		Natural Spline		N
	RR	95% CI	RR	95% CI	RR	95% CI	
CVD[b]							
PM_{10}	1.015	0.987,1.043	1.015	0.988,1.044	1.025	0.995,1.057	1081
O_3	1.048	1.026,1.071	1.053	1.030,1.076	1.062	1.033,1.092	3196
BS	1.032	1.016,1.048	1.031	1.015,1.047	1.029	1.010,1.048	3267
SO_2	1.039	1.022,1.056	1.037	1.020,1.054	1.038	1.018,1.057	3191
NO_2	1.028	1.015,1.041	1.025	1.012,1.038	1.021	1.005,1.036	3177
CO	1.044	1.012,1.077	1.041	1.009,1.074	1.046	1.010,1.083	2153
SO_4^{2-}	1.021	0.981,1.063	1.022	0.982,1.064	1.027	0.984,1.072	876
NO_3^-	1.024	0.983,1.066	1.025	0.984,1.068	1.035	0.990,1.083	876
Sec[c]	1.026	0.982,1.072	1.028	0.984,1.073	1.036	0.988,1.086	876
COPD[d]							
PM_{10}	1.096	1.014,1.185	1.099	1.017,1.188	1.097	1.007,1.195	1081
O_3	0.994	0.932,1.059	1.009	0.947,1.075	1.023	0.940,1.113	3196
BS	1.072	1.026,1.120	1.070	1.024,1.117	1.072	1.018,1.129	3267
SO_2	1.056	1.007,1.107	1.049	1.001,1.100	1.051	0.996,1.109	3191
NO_2	1.090	1.052,1.129	1.077	1.040,1.116	1.081	1.036,1.129	3177
CO	1.194	1.099,1.298	1.184	1.090,1.287	1.190	1.085,1.305	2153
SO_4^{2-}	1.115	0.996,1.248	1.119	1.000,1.252	1.118	0.992,1.261	876
NO_3^-	1.077	0.958,1.211	1.081	0.961,1.215	1.098	0.966,1.247	876
Sec	1.115	0.984,1.262	1.119	0.989,1.268	1.129	0.987,1.291	876
Pneumonia							
PM_{10}	1.167	1.058,1.287	1.169	1.060,1.289	1.176	1.057,1.309	1081
O_3	1.130	1.049,1.217	1.147	1.065,1.235	1.150	1.048,1.263	3196
BS	1.126	1.064,1.192	1.122	1.060,1.188	1.137	1.063,1.215	3267
SO_2	1.102	1.034,1.174	1.092	1.025,1.164	1.093	1.016,1.175	3191
NO_2	1.146	1.094,1.201	1.134	1.082,1.188	1.153	1.089,1.221	3177
CO	1.276	1.143,1.426	1.266	1.133,1.414	1.257	1.112,1.423	2153
SO_4^{2-}	1.098	0.954,1.264	1.106	0.961,1.273	1.104	0.948,1.284	876
NO_3^-	1.202	1.040,1.389	1.208	1.045,1.396	1.251	1.068,1.466	876
Sec	1.175	1.006,1.372	1.184	1.014,1.382	1.204	1.018,1.424	876

[a] Averages of lag 0–6 days, except ozone, sulfate, nitrate and sec (lag 1).

[b] CVD = Cardiovascular disease.

[c] Sec = sum of SO_4^{2-} and NO_3^- as an approximation of secondary aerosol.

[d] COPD = Chronic obstructive pulmonary disease.

model (Dominici et al 2002). Thus an additional LOESS term for influenza was added to the model. This did not alter effect estimates nor the sensitivity to different models for date (data not shown). Next, we replaced the bilinear temperature variables with two LOESS functions with a span of 0.7 (3 degrees of freedom) for the warm and cold continuous temperature variables. Specification of LOESS functions for temperature changed the effect estimates of the original model marginally (Table 6). Effect estimates increased slightly. When the LOESS functions were replaced with natural splines, somewhat larger changes of the air pollution effect estimates occurred: increases for ozone and decreases for the other pollutants. Changes were larger for the weekly average than for the previous day concentration. Most effect estimates remained significant however.

Season-specific analyses with the two LOESS functions for temperature in the model did not show any systematic difference between LOESS models and spline models (for date and temperature) (Table 7).

Table 4. Sensitivity of Specific Cardiovascular Mortality Associations with Air Pollution to Confounder Model Specification

	Pollutant[a]	GAM Default		GAM Strict		Natural Spline		
		RR	95% CI	RR	95% CI	RR	95% CI	N
CVD[b]	PM_{10}	1.015	0.987,1.043	1.015	0.988,1.044	1.025	0.995,1.057	1081
	O_3	1.048	1.026,1.071	1.053	1.030,1.076	1.062	1.033,1.092	3196
	BS	1.032	1.016,1.048	1.031	1.015,1.047	1.029	1.010,1.048	3267
	SO_2	1.039	1.022,1.056	1.037	1.020,1.054	1.038	1.018,1.057	3191
	NO_2	1.028	1.015,1.041	1.025	1.012,1.038	1.021	1.005,1.036	3177
	CO	1.044	1.012,1.077	1.041	1.009,1.074	1.046	1.010,1.083	2153
MI[c]	PM_{10}	1.005	0.964,1.048	1.006	0.965,1.049	1.016	0.970,1.063	1081
	O_3	1.026	0.994,1.060	1.034	1.001,1.067	1.043	1.001,1.086	3196
	BS	1.017	0.994,1.041	1.015	0.992,1.039	1.009	0.982,1.037	3267
	SO_2	1.015	0.991,1.039	1.011	0.987,1.035	1.011	0.984,1.039	3191
	NO_2	1.017	0.998,1.035	1.011	0.992,1.030	1.002	0.980,1.025	3177
	CO	1.050	1.004,1.099	1.045	0.999,1.094	1.043	0.991,1.097	2153
RYTM[d]	PM_{10}	1.041	0.932,1.163	1.044	0.934,1.166	1.041	0.922,1.175	1081
	O_3	1.074	0.977,1.181	1.092	0.993,1.200	1.114	0.988,1.255	3196
	BS	1.071	1.001,1.146	1.068	0.998,1.143	1.058	0.977,1.145	3267
	SO_2	1.046	0.971,1.128	1.040	0.964,1.122	1.032	0.947,1.126	3191
	NO_2	1.030	0.976,1.088	1.020	0.966,1.076	1.013	0.948,1.082	3177
	CO	1.062	0.937,1.203	1.053	0.929,1.194	1.031	0.898,1.184	2153
FAIL[e]	PM_{10}	1.036	0.960,1.118	1.039	0.963,1.121	1.047	0.963,1.139	1081
	O_3	1.079	1.009,1.154	1.090	1.019,1.165	1.102	1.012,1.199	3196
	BS	1.081	1.031,1.134	1.079	1.029,1.132	1.083	1.024,1.146	3267
	SO_2	1.098	1.043,1.156	1.093	1.038,1.151	1.100	1.037,1.167	3191
	NO_2	1.064	1.024,1.106	1.059	1.019,1.101	1.059	1.011,1.110	3177
	CO	1.109	1.012,1.216	1.102	1.005,1.208	1.138	1.027,1.260	2153
CERE[f]	PM_{10}	1.031	0.971,1.094	1.032	0.973,1.096	1.038	0.973,1.108	1081
	O_3	1.071	1.023,1.122	1.082	1.033,1.133	1.091	1.029,1.157	3196
	BS	1.041	1.007,1.077	1.040	1.006,1.075	1.036	0.996,1.077	3267
	SO_2	1.066	1.029,1.104	1.062	1.026,1.100	1.062	1.020,1.106	3191
	NO_2	1.047	1.020,1.076	1.041	1.014,1.069	1.040	1.006,1.074	3177
	CO	1.048	0.982,1.118	1.043	0.977,1.113	1.062	0.989,1.142	2153
TROM[g]	PM_{10}	1.010	0.894,1.143	1.011	0.894,1.144	1.015	0.887,1.161	1081
	O_3	1.140	1.032,1.259	1.155	1.046,1.276	1.166	1.028,1.322	3196
	BS	1.044	0.972,1.122	1.043	0.971,1.121	1.038	0.954,1.130	3267
	SO_2	1.126	1.044,1.213	1.125	1.044,1.213	1.137	1.043,1.239	3191
	NO_2	1.002	0.946,1.062	0.997	0.941,1.056	0.991	0.924,1.064	3177
	CO	1.065	0.926,1.224	1.058	0.920,1.217	1.089	0.933,1.271	2153

[a] Average of lag 0–6 days, except ozone (lag 1).

[b] CVD = all cardiovascular disease combined.

[c] MI = myocardial infarction.

[d] RYTM = arrhythmia.

[e] FAIL = heart failure.

[f] CERE = cerebrovascular disease.

[g] TROM = thrombosis related diseases.

Table 5. Sensitivity of Two Pollutant Models for Total Mortality

Model	Pollutant[a]	GAM Default		GAM Strict		Spline		
		RR	95% CI	RR	95% CI	RR	95% CI	N
$PM_{10}+O_3$	PM_{10}	1.025	1.007,1.043	1.026	1.007,1.044	1.021	1.001,1.042	1081
	O_3	1.045	1.018,1.074	1.050	1.022,1.079	1.048	1.014,1.083	1081
$PM_{10}+SO_2$	PM_{10}	0.983	0.960,1.006	0.987	0.964,1.011	0.986	0.959,1.014	1081
	SO_2	1.137	1.077,1.201	1.123	1.063,1.186	1.115	1.045,1.189	1081
$PM_{10}+NO_2$	PM_{10}	0.981	0.957,1.006	0.988	0.964,1.013	0.989	0.958,1.020	1081
	NO_2	1.052	1.028,1.076	1.044	1.020,1.068	1.037	1.007,1.067	1081
$PM_{10}+CO$	PM_{10}	1.038	1.005,1.072	1.044	1.011,1.079	1.056	1.015,1.099	1081
	CO	0.969	0.914,1.028	0.959	0.904,1.016	0.927	0.864,0.996	1081
$BS+O_3$	BS	1.037	1.026,1.048	1.034	1.023,1.045	1.029	1.016,1.043	3196
	O_3	1.041	1.026,1.056	1.046	1.031,1.061	1.048	1.029,1.067	3196
$BS+SO_2$	BS	1.013	0.999,1.028	1.012	0.998,1.026	1.014	0.995,1.034	3191
	SO_2	1.027	1.013,1.042	1.025	1.010,1.040	1.015	0.996,1.034	3191
$BS+NO_2$	BS	1.009	0.993,1.025	1.012	0.996,1.028	1.016	0.995,1.037	3177
	NO_2	1.026	1.014,1.038	1.019	1.007,1.032	1.010	0.993,1.027	3177
$BS+CO$	BS	1.048	1.015,1.081	1.050	1.018,1.084	1.062	1.021,1.104	2153
	CO	0.980	0.933,1.030	0.972	0.925,1.021	0.958	0.902,1.017	2153
$PM_{10}+BS$	PM_{10}	1.009	0.974,1.046	1.015	0.980,1.052	1.028	0.977,1.082	1081
	BS	1.017	0.977,1.059	1.010	0.970,1.052	0.988	0.928,1.051	1081
$SO_4^{2-}+PM_{10}$	SO_4^{2-}	1.029	0.997,1.062	1.029	0.997,1.062	1.028	0.994,1.064	876
	PM_{10}	1.005	0.981,1.031	1.007	0.983,1.033	1.003	0.974,1.033	876
$SO_4^{2-}+BS$	SO_4^{2-}	1.027	0.997,1.058	1.028	0.998,1.059	1.029	0.996,1.063	876
	BS	1.012	0.985,1.041	1.012	0.985,1.041	1.004	0.969,1.039	876
$SO_4^{2-}+O_3$	SO_4^{2-}	1.037	1.010,1.064	1.038	1.012,1.065	1.036	1.008,1.066	876
	O_3	1.070	1.038,1.104	1.075	1.043,1.108	1.076	1.035,1.118	876
$SO_4^{2-}+SO_2$	SO_4^{2-}	1.017	0.990,1.045	1.019	0.992,1.047	1.015	0.986,1.046	876
	SO_2	1.100	1.047,1.156	1.093	1.040,1.148	1.097	1.037,1.161	876
$SO_4^{2-}+NO_2$	SO_4^{2-}	1.019	0.991,1.048	1.022	0.994,1.051	1.019	0.989,1.051	876
	NO_2	1.027	1.008,1.047	1.023	1.004,1.043	1.020	0.997,1.044	876
$SO_4^{2-}+CO$	SO_4^{2-}	1.035	1.005,1.066	1.037	1.007,1.068	1.040	1.007,1.073	876
	CO	0.990	0.951,1.030	0.988	0.949,1.028	0.971	0.927,1.017	876
$NO_3^-+PM_{10}$	NO_3^-	1.040	1.007,1.074	1.040	1.007,1.074	1.042	1.007,1.078	876
	PM_{10}	1.002	0.977,1.027	1.004	0.979,1.029	0.998	0.970,1.027	876
NO_3^-+BS	NO_3^-	1.037	1.007,1.069	1.038	1.008,1.070	1.041	1.008,1.076	876
	BS	1.008	0.981,1.037	1.009	0.981,1.037	0.999	0.964,1.034	876
$NO_3^-+O_3$	NO_3^-	1.048	1.021,1.076	1.049	1.022,1.077	1.050	1.019,1.081	876
	O_3	1.074	1.041,1.107	1.078	1.046,1.111	1.080	1.039,1.123	876
$NO_3^-+SO_2$	NO_3^-	1.025	0.997,1.055	1.027	0.999,1.057	1.024	0.993,1.057	876
	SO_2	1.095	1.042,1.151	1.087	1.035,1.143	1.090	1.029,1.155	876
$NO_3^-+NO_2$	NO_3^-	1.026	0.995,1.058	1.029	0.999,1.061	1.030	0.996,1.065	876
	NO_2	1.023	1.003,1.045	1.019	0.999,1.040	1.016	0.991,1.041	876
$NO_3^- +CO$	NO_3^-	1.046	1.015,1.078	1.048	1.017,1.079	1.053	1.019,1.088	876
	CO	0.984	0.944,1.025	0.982	0.943,1.023	0.964	0.921,1.009	876

[a] Average of lag 0–6 days, except for ozone, nitrate, and sulfate (lag 1).

Table 6. Sensitivity of Total Mortality Analyses for Substituting Original Bilinear Temperature Terms by LOESS Functions of Temperature

Pollutant	GAM Default		GAM Strict		Natural Spline		N
	RR	95% CI	RR	95% CI	RR	95% CI	
PM_{10} lag1	1.020	1.005,1.036	1.020	1.004,1.035	1.015	0.998,1.033	1092
PM_{10} avg	1.028	1.010,1.047	1.027	1.008,1.045	1.012	0.989,1.036	1081
O_3 lag1	1.033	1.019,1.048	1.039	1.024,1.053	1.043	1.024,1.062	3196
O_3 avg	1.018	1.003,1.033	1.042	1.027,1.057	1.060	1.032,1.089	3191
BS lag1	1.022	1.013,1.032	1.021	1.011,1.030	1.019	1.008,1.030	3267
BS avg	1.034	1.023,1.044	1.029	1.019,1.040	1.022	1.008,1.036	3267
SO_2 lag1	1.029	1.019,1.039	1.027	1.017,1.037	1.023	1.012,1.035	3196
SO_2 avg	1.044	1.032,1.055	1.038	1.027,1.049	1.028	1.013,1.043	3191
NO_2 lag1	1.031	1.022,1.040	1.029	1.020,1.038	1.025	1.015,1.035	3184
NO_2 avg	1.035	1.027,1.044	1.030	1.021,1.038	1.020	1.009,1.032	3177
CO lag1	1.038	1.021,1.055	1.035	1.018,1.052	1.028	1.008,1.047	2158
CO avg	1.055	1.034,1.077	1.047	1.026,1.068	1.039	1.012,1.067	2153
SO_4^{2-} lag1	1.033	1.007,1.061	1.033	1.007,1.060	1.026	0.997,1.056	876
SO_4^{2-} avg	1.021	0.997,1.045	1.021	0.998,1.046	1.008	0.978,1.038	871
NO_3^- lag1	1.044	1.017,1.072	1.044	1.017,1.072	1.037	1.007,1.069	876
NO_3^- avg	1.033	1.009,1.057	1.031	1.008,1.056	1.017	0.987,1.047	871
Sec[a] lag1	1.046	1.016,1.076	1.045	1.016,1.075	1.038	1.005,1.071	876
Sec[a] avg	1.030	1.005,1.056	1.029	1.004,1.055	1.014	0.982,1.046	871

[a] Sec = sum of SO_4^{2-} and NO_3^- as an approximation of secondary aerosol.

Table 7. Sensitivity of Season Specific Analysis to Different Models

Pollutant[a]	Season	GAM Default		GAM Strict		Natural Spline		N
		RR	95% CI	RR	95% CI	RR	95% CI	
PM_{10}	winter	1.022	1.000,1.045	1.023	1.001,1.046	1.028	1.001,1.057	439
	summer	1.090	1.034,1.148	1.088	1.033,1.147	1.072	0.998,1.151	459
O_3	winter	0.988	0.960,1.015	0.987	0.960,1.015	1.008	0.970,1.046	1271
	summer	1.074	1.053,1.097	1.073	1.052,1.095	1.071	1.047,1.096	1377
BS	winter	1.025	1.013,1.037	1.024	1.012,1.037	1.023	1.007,1.039	1341
	summer	1.111	1.067,1.157	1.108	1.064,1.154	1.141	1.065,1.222	1377
SO_2	winter	1.033	1.020,1.046	1.031	1.018,1.044	1.035	1.020,1.049	1271
	summer	1.068	1.021,1.118	1.057	1.011,1.106	1.058	0.995,1.124	1377
NO_2	winter	1.025	1.014,1.036	1.024	1.013,1.035	1.021	1.006,1.035	1271
	summer	1.042	1.023,1.061	1.039	1.020,1.058	1.049	1.021,1.077	1377
CO	winter	1.038	1.013,1.063	1.037	1.013,1.062	1.043	1.013,1.074	869
	summer	1.199	1.108,1.296	1.188	1.099,1.284	1.231	1.105,1.371	918

[a] Average of lag 0–6 days, except for ozone (lag 1).
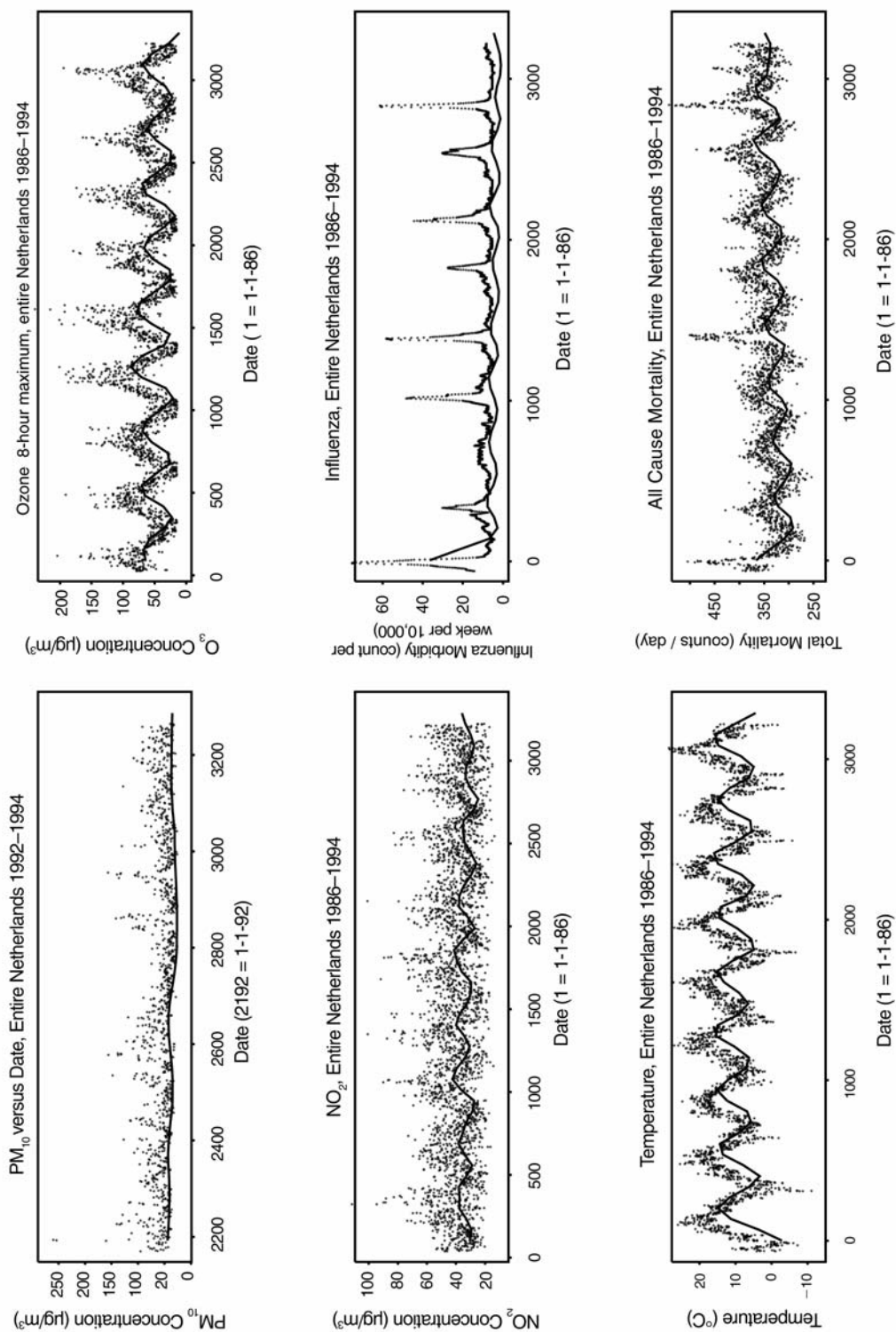
Figure 2. Temporal pattern of air pollution, mortality, and weather.

DISCUSSION

Specification of stricter convergence criteria for the original GAM resulted in very similar RR estimates for air pollution. Air pollution effect estimates were in general very similar between GAM and natural spline models. Standard errors of the natural spline model were in general between 10% and 25% higher.

The small change in effect estimates is not due to a lack of seasonal variation: mortality peaks in the winter season as do all air pollutants (including PM_{10}, sulfate and nitrate) except ozone (Figure 2). The small differences between models might be due to the use of only one LOESS function in the original model (Dominici et al 2002). Using an additional LOESS function for influenza did not make any difference for the air pollution effect estimates. Thus, just the number of LOESS terms is not related to differences between spline and GAMs. When we added LOESS functions for temperature to the model, the difference between LOESS and natural spline models increased. This is consistent with observations by Ramsay and colleagues (2003), who have stated that GAMs do not perform well in cases of concurvity between pollution and other confounders. Concurvity is the nonparametric analog of multicollinearity. This interpretation is supported by the larger changes in effect estimates that we observed for weekly average concentrations compared to single day lags and the complete lack of sensitivity to model choice in season specific analyses. Alternatively, it is possible that the natural spline model does not capture the seasonal variation sufficiently. In this dataset, increasing the degrees of freedom for the natural spline model only slightly increased the effect estimates. Further research is necessary to determine which of the two models gives the least biased results.

Overall, the conclusion of these reanalyses is that the specification of stricter convergence criteria and the use of natural splines instead of GAM did not change any of the conclusions drawn in the two original papers.

REFERENCES*

Dominici F, Mc Dermott A, Zeger SL, Samet JM. 2002. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiol 156:193–203.

Hoek G, Brunekreef B, Fischer P, van Wijnen J. 2001. The association between air pollution and heart failure, arrhythmia, embolism, thrombosis, and other cardiovascular causes of death in a time series study. Epidemiology 12:355–357.

Hoek G, Brunekreef B, Verhoeff A, van Wijnen J, Fischer P. 2000. Daily mortality and air pollution in The Netherlands. J Air Waste Manage Assoc 50:1380–1389.

Hoek G, Verhoeff A, Fischer P. 1997. Daily Mortality and Air Pollution in The Netherlands, 1986–1994. Report 1997-481. Department of Environmental Health, Wageningen University, Wageningen, The Netherlands.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

ABBREVIATIONS AND OTHER TERMS

BS	black smoke
CI	confidence interval
CO	carbon monoxide
CVD	cardiovascular disease
GAM	generalized additive model
LOESS	locally weighted smoothers
NO_2	nitrogen dioxide
NO_3^-	nitrate
ns	natural spline
O_3	ozone
PM_{10}	particulate matter of 10 µm mass median aerodynamic diameter
RR	relative risk
SO_2	sulfur dioxide
SO_4^{2-}	sulfate

* Bold type identifies publications containing the original analyses revised in this short communication report.

Associations of Particulate Matter Components with Daily Mortality and Morbidity in Detroit, Michigan

Kazuhiko Ito

ABSTRACT

To examine the influence of the generalized additive model (GAM*) convergence problem on the results of the original Detroit data analysis, the associations between particulate matter (PM) components and daily mortality/morbidity were reexamined using stringent convergence criteria as well as generalized linear models (GLMs) that approximated the original GAMs. Generally, both the GAMs with stringent convergence criteria and GLMs resulted in smaller estimated relative risks (RRs) than those reported in the original study. The reductions in the estimated RRs did not differ across the PM indices. Thus, the conclusions of the original study regarding the lack of relative role of PM by size and chemical characteristics remain the same. Sensitivity analyses using various weather models and degrees of freedom (*df*) of time trend smoothing indicate that the extent of difference in RR estimates across different model specifications can be larger than that caused by the GAM convergence problem.

INTRODUCTION

Lippmann and colleagues (2000) examined associations between PM components and daily mortality and elderly hospital admission data in Detroit, Michigan. The key feature of the study was that multiple PM components were available from a monitor in Windsor, Ontario, that was only a few miles from downtown Detroit. These PM indices included total suspended particles (TSP), PM less than 10 μm in diameter (PM_{10}), and sulfate from TSP filters for the 1985–1990 study period (mortality analysis

only) and PM_{10}, $PM_{2.5}$ (PM < 2.5 μm in diameter), $PM_{10-2.5}$ (PM_{10} minus $PM_{2.5}$), sulfate (SO_4^{2-}), and H^+ for the 1992–1994 study period (for which both mortality and elderly hospital admissions were analyzed). Gaseous pollutants (sulfur dioxide [SO_2], nitrogen dioxide [NO_2], ozone [O_3], and carbon monoxide [CO]) were also available for these study periods.

The study addressed a specific hypothesis regarding the relative importance of PM components: that the closest associations are attributable to PM, more specifically to PM_{10} and/or $PM_{2.5}$ and especially the SO_4^{2-} and H^+ components of $PM_{2.5}$. Two major findings from the original study were as follows:

1. Our results were generally inconsistent with our study hypothesis, particularly regarding the influence of aerosol H^+ components of $PM_{2.5}$, that the relative particle metric effect size and strength of associations with mortality and morbidity outcomes would be as follows: $H^+ > SO_4^{2-} > PM_{2.5} > PM_{10} > TSP$.

2. Generally, the PM mass indices were more significantly associated with health outcomes than were H^+ or SO_4^{2-}. When both H^+ and SO_4^{2-} were significantly associated, SO_4^{2-} was associated even more strongly with the outcomes. These results suggest that PM components other than H^+, including the coarse component of PM_{10} (ie, $PM_{10-2.5}$), may be harmful.

Recently, Dominici and coworkers (2002) reported that using the GAM with default convergence criteria in S-Plus statistical software could lead to biased results when the coefficients are small and smoothing terms are included. To examine the influence of this convergence problem on the results of the original study, the associations between PM components and daily mortality and morbidity were reexamined using stringent convergence criteria as well as GLMs that approximated the original GAMs. Additionally, several sensitivity analyses using alternative model specifications were conducted.

METHODS

This revised analysis consisted of (1) revised analysis of the original data using the exact final GAM Poisson model specifications run in the original study but using more stringent convergence criteria; (2) analysis of the original data using a GLM Poisson model approximating the original GAM Poisson model; and (3) sensitivity analyses using alternative model specifications. The original model development is first briefly summarized, and then the model specification for reanalyses and additional analyses are described.

The original study (Lippmann et al 2000) involved sequential development of health-effects models in stages. First, adjustments for confounding time trends were considered by evaluating the fitted health outcome series, residual dispersion, and autocorrelation in Poisson GAM with smoothing splines (as implemented in S-Plus as s( )) with several candidate degrees of freedom. Noting that the apparent influenza epidemics were captured and the residual overdispersion and autocorrelation were minimized (but not underdispersed or negatively autocorrelated), the degree of freedom that was equivalent to changing directions in a period of approximately one month (ie, 12 *df* per year) was chosen for all the outcomes for both the 1985–1990 and 1992–1994 study periods. The sole exception was respiratory mortality for the 1985–1990 study period, for which a three-month period (ie, 4 *df* per year) was chosen to avoid underdispersion and negative autocorrelation caused by a greater degree of freedom. Second, using the degree of freedom chosen above, several alternative weather model specifications were evaluated for their consistency with biological plausibility and model fit (ie, Akaike information criterion [AIC]). The final weather model chosen contained a locally weighted smoother (LOESS) of same-day temperature (to model warm temperature effect) with a span of 0.5, a LOESS of the average of temperatures lagged 1 through 3 days (to model cold temperature effects) with a span of 0.5, and a hot-and-humid indicator variable (mean temperature above 80°F and relative humidity above 70%; to model hot and humid interaction), in addition to the seasonal smoothing term described above and day-of-week indicator variables. Third, the lagged (0-day lag through 3-day lag and multiday average thereof) pollution variables were added to the Poisson model with the health outcome variables fitted to season and weather. That is, the logarithm of the fitted values from Poisson GAMs with the temporal trends and weather models were included in the second Poisson model with lagged pollution variables. This inclusion implies that the second-stage Poisson model that estimated pollution

coefficients was a GLM, not a GAM, since no smoothing was done. Since the Poisson model to estimate the time trends and weather effects was modeled using the GAM Poisson model with lax default convergence criteria, however, the fitted time trend/weather models still fit incompletely, which affected the corresponding pollution effect estimates.

The selection of the weather model described above in part considered AIC, although, as mentioned in the original report, AIC was not as important as consistency with biological plausibility. Since using the lax default convergence criteria can lead to an underachievement of possible fit, the deviance is generally larger for the GAM Poisson model with default convergence criteria than for the model with more stringent convergence criteria. Furthermore, the extent to which the deviance is reduced for the GAM with more stringent convergence criteria can vary across the alternative weather models. Therefore, GAM Poisson models with more stringent convergence criteria could have resulted in different rankings of AIC values across the alternative weather models evaluated in the original analysis. However, the reanalysis was conducted using exactly the same GAM specifications as were used in the original study because a comparison of the results using different weather model specifications would obscure an examination of the effects of the convergence criteria problem. The effects of using alternative model specifications were examined as part of the sensitivity analysis.

A revised analysis of GAM Poisson models was conducted using more stringent convergence criteria as suggested by Dominici and coworkers (2002): the convergence precision (epsilon) was set to 10^{-14} and maximum iteration was set to 1000 for both the local scoring and backfitting algorithms.

The GLM specification approximated the original GAMs. Natural splines (ns( ) as implemented in S-Plus) were used as smoothing terms. To model time trend, the same degree of freedom was used as in the smoothing splines in the GAMs described above, with the default placement of knots. For weather models, to approximate LOESS smoothing with a span of 0.5 in the GAM, natural splines with 2 *df* were used.

Since the two-step approach used in the original analysis could have prevented the pollution indices from competing fairly with weather and temporal trends, analyses were repeated using the same GAMs (with stringent convergence criteria) with all variables included simultaneously. Similarly, analyses were repeated using comparable GLMs with natural splines.

While the weather and temporal trend model specification used in the original analysis was developed with considerations that were specific to this data set, a

question remains as to how sensitive the results could have been had we chosen different weather model specifications or different degrees of freedom for time trends. To address this issue, original weather specification, two additional weather models, and no weather model were each applied to a limited set of pollution variables (1-day lagged PM_{10}) and health outcomes (total mortality and hospital admissions for pneumonia) using GLM with natural splines. For each of the four weather specifications, five levels of degrees of freedom (corresponding approximately to one year, six months, three months, one month, and two weeks) were applied for natural splines. The two additional weather specifications were similar to those in published studies, except that natural splines were used instead of LOESS or smoothing splines. One model was similar to that used in the Harvard Six-Cities time-series study (Schwartz et al 1996) in which LOESS terms of same-day temperature and same-day dewpoint with spans of 0.5 were included. To approximate this model, natural splines were used with 2 *df* instead of LOESS terms. Another weather model evaluated was similar to that used in the analysis of 90 US cities by Samet and associates (2000a,b), in which smoothing splines of the same-day temperature ($df = 6$), the average of temperatures lagged 1 through 3 days ($df = 6$), same-day dewpoint ($df = 3$), and the average of dewpoints lagged 1 through 3 days ($df = 3$) were included. Instead of smoothing splines, natural splines were used in this sensitivity analysis.

Combinations of four health outcomes (total mortality from all causes [excluding injury and poisoning], circulatory diseases, respiratory disorders, and other mortality [total mortality excluding injury, poisoning, circulatory, and respiratory causes]), eight pollutants (TSP, PM_{10}, TSP-PM_{10}, sulfate from TSP filters, O_3, SO_2, NO_2, and CO), and ten lag/average days (0, 1, 2, 3, averages of 0–1, 1–2, 2–3, 0–2, 1–3, and 0–3) were examined for the 1985–1990 study period. Likewise, ten outcomes (four mortality series described above plus hospital admissions for pneumonia, chronic obstructive pulmonary disease [COPD], ischemic heart disease, dysrhythmias, heart failure, and stroke in the elderly), nine pollutants (PM_{10}, $PM_{2.5}$, $PM_{10-2.5}$, SO_4^{2-}, H^+, O_3, NO_2, SO_2, and CO) and the ten lag/average days were examined for the 1992–1994 study period. Thus, 1220 (= 4 × 8 × 10 + 10 × 9 × 10) RR estimates were calculated for each set of analysis for single pollutant models. In the original report, all of these 1220 RR estimates and their 95% confidence intervals (CIs) were presented as appendices. In addition, in the original report, tables (PM indices only) of RRs and 95% CIs for lags with the most significant estimates and corresponding figures (both PM indices and gaseous pollutants) were presented. In this reanalysis, tables for PM indices (only single-pollutant models) and corresponding

figures with original results, revised results from analyses with stringent convergence criteria, and GLM results are presented. In addition, distributions of the resulting 1220 RRs are described to facilitate a more comprehensive evaluation of these results.

RESULTS

Figure 1 shows a scatter plot of the 1220 RRs per 5th-to-95th percentile increment of the pollution index distribution for GAM with default convergence criteria (ie, the original study) versus GAM with stringent convergence criteria for all the health outcomes, pollutants, and lag/averaging days combinations. For a large fraction, the GAM with stringent convergence criteria provided smaller RR estimates than the GAM with default convergence criteria. The difference in RR estimates appears to be independent of the level of RR estimates. The median difference in RR per 5th-to-95th percentile increment was 0.0071. Figure 2 shows a scatter plot of the GLM results versus GAM with stringent convergence criteria for all the health outcomes, pollutants, and lag/averaging days combinations. Overall, the points generally scattered around a curve with a slope of 1. How-
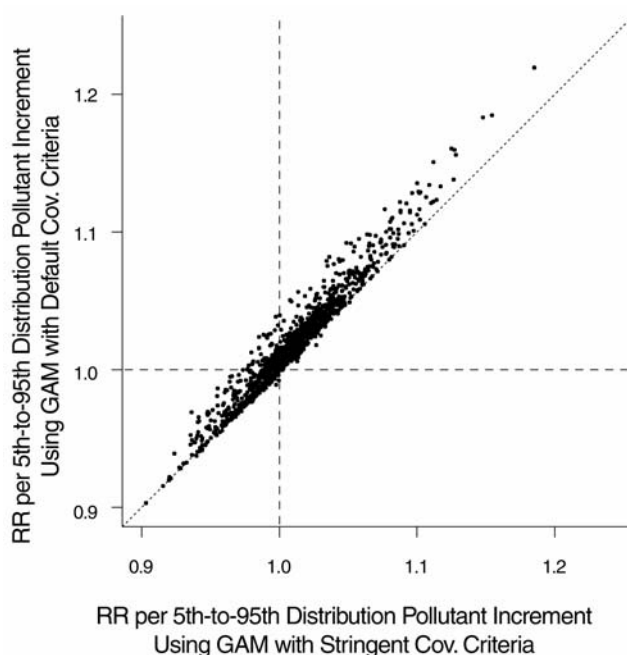
Figure 1. Plot of the 1220 relative risks (RRs) obtained with the original (default) and revised (stringent) GAMs for all health outcomes per 5th-to-95th percentile increment in pollution index distribution for all pollutants and all lags and averaging-day combinations.

145
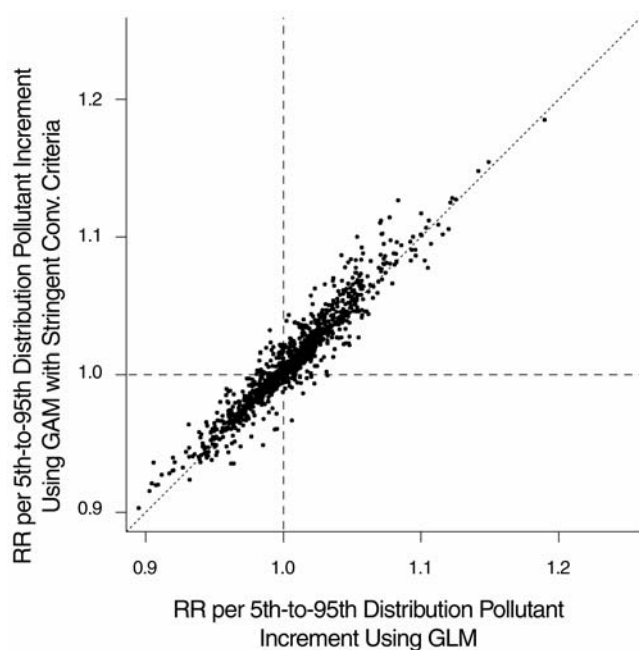
Figure 2. Plot of the 1220 relative risks (RRs) obtained with the revised (stringent) GAM and the GLM with natural splines for all health outcomes per 5th-to-95th percentile increment in pollution index distribution for all pollutants and all lags and averaging-day combinations.

ever, the median difference in RR was 0.0028; the RR estimates were smaller for the GLM results.

In the original report, the RRs for all the lag/averaging periods were presented in the appendix, but the main text focused on a presentation of tables and figures comparing RRs across PM indices and gaseous pollutants at their most significant lags. While such an approach may be useful to address the original study hypothesis, a more comprehensive summary requires an examination of the distribution of the RRs by pollutant. Figure 3 shows box plots of all 1220 RRs by pollutant for (a) the original analysis and (b) the revised analysis with stringent convergence criteria. Obviously, the RRs in these distributions are not independent events, because RRs for all the lag/averages and health outcomes were included. These distributions also included RRs for a mixture of health outcomes, including those that were not associated with any PM indices (ie, hospital admissions for stroke and dysrhythmias). With these limitations in mind, Figure 3 provides an overall comparison of the distributions of RR estimates by pollutant. The distributions of RRs were lower for the GAM with stringent convergence criteria than for the GAM with default convergence criteria across all pollutants. However, the relative extent of RR estimates across pollutants in the revised analysis remains the same as that in the original study. There appear to have been no differential reductions in RR

estimates across the PM indices or gaseous pollutants. The reductions in RR estimates for SO_2, NO_2, and CO were such that their medians for all the lag/averages and outcomes were essentially 1 (no effects).

Tables 1 through 12 show results for the original model (GAM with default convergence criteria), GAM with stringent convergence criteria, and GLM approximating the original GAMs for PM indices at lags that were the most significant in the original analyses. These tables also show the corresponding table numbers in the original report. To facilitate comparison, the same information (except for hospital admissions for dysrhythmias and stroke) is presented in Figures 4 through 6. Again, it can be seen that, in the majority of cases, reductions in RR estimates did not differ across the PM indices. Also, the tables and figures show that the widths of the 95% CIs for these three models were essentially the same as expected; while the smoothing terms for the seasonal trend and weather models were fitted in GAMs or GLMs in the first stage of the regression, the coefficients for pollutants in the second stage were always fitted using GLM, as described in the Methods. Except for a few instances in which the original results had lags with almost the same RR estimates, the revised analysis did not change the lag structure of the associations.

Figure 7 (corresponding to Figure 17 in the original report) shows the results of GAMs, original and using stringent convergence criteria, in which $PM_{2.5}$ and $PM_{10-2.5}$ were included simultaneously. As can be seen, the pattern found in the original analysis was essentially unchanged.

Figure 8 shows box plots of all 1220 RRs by pollutant for GAMs (with stringent convergence criteria) in which weather, smooth time trend, and pollution variables were included simultaneously, as well as plots for the corresponding GLMs. The relative magnitude of the RRs across pollutants was similar to those in the original analysis or reanalysis. The levels of RRs in these results were larger than those in the original analysis or reanalysis (compare with Figure 3), however. For example, the median PM_{10} RRs in Figure 3(b) is 1.017, whereas it is 1.026 in Figure 8(a). As expected, the standard errors of coefficients in Figure 8(b), showing GLM results, were larger (median 16%) than those in Figure 8(a).

The sensitivity of PM_{10} coefficients (at lag 1 day) to the weather model specification and the degree of freedom for temporal trends in the 1992–1994 study period is shown in Figure 9 (for total mortality) and Figure 10 (for pneumonia admissions among the elderly). For total mortality with degree of freedom corresponding to less than three months, the coefficients for the three weather models were relatively stable but ranged from 0.00051 to 0.00081. For pneumonia
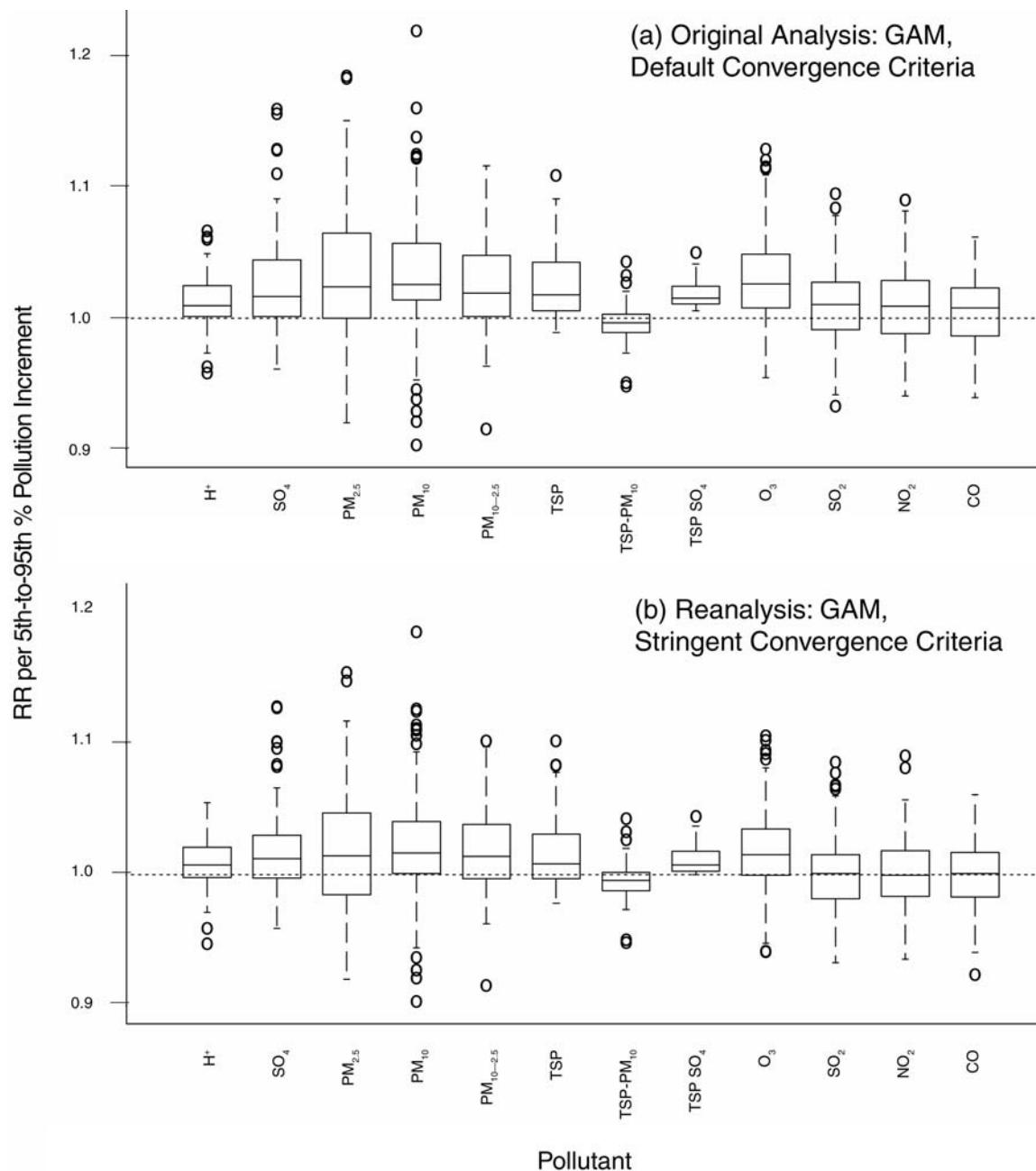
Figure 3. Box plots of all 1220 relative risks (RRs) by pollutant for (a) GAM with default convergence criteria and (b) revised GAM with stringent convergence criteria.

admissions, the coefficients were more sensitive to the degree of freedom. For degree of freedom corresponding to less than three months, the coefficients for the three weather models ranged between 0.00075 to 0.00293. The coefficients for the model without weather adjustment were up to two to three times larger than those with weather adjustment for both mortality and pneumonia admissions, but ranged from 0.00051 to 0.00081. For pneumonia admissions, the

coefficients were more sensitive to the degree of freedom. For degree of freedom corresponding to less than three months, the coefficients for the three weather models ranged between 0.00075 to 0.00293. The coefficients for the model without weather adjustment were up to two to three times larger than those with weather adjustment for both mortality and pneumonia admissions.

Table 1. Relative Risks (95% CI below) for Total Mortality for PM Indices per 5th-to-95th Percentile Increment, 1985–1990[a]

	Original GAM (default)	GAM (stringent)	GLM
PM_{10} (lag 1)	1.026 (1.003,1.049)	1.014 (0.992,1.037)	1.010 (0.988,1.033)
TSP (lag 3)	1.018 (0.996,1.040)	1.009 (0.987,1.031)	1.003 (0.982,1.025)
TSP-PM_{10} (lag 1)	1.001 (0.978,1.025)	1.000 (0.977,1.024)	0.998 (0.975,1.022)
TSP sulfates (lag 3)	1.015 (0.995,1.035)	1.008 (0.988,1.028)	1.005 (0.986,1.026

[a] Corresponds to Table 5 in Lippmann and colleagues 2000, p 18.

Table 3. Relative Risks (95% CI below) for Mortality from Respiratory Causes for PM Indices per 5th-to-95th Percentile Increment, 1985–1990[a]

	Original GAM (default)	GAM (stringent)	GLM
PM_{10} (lag 1)	1.123 (1.036,1.218)	1.114 (1.027,1.208)	1.077 (0.994,1.168)
TSP (lag 1)	1.109 (1.028,1.197)	1.102 (1.021,1.190)	1.071 (0.993,1.156)
TSP-PM_{10} (lag 1)	1.043 (0.960,1.133)	1.043 (0.959,1.134)	1.035 (0.952,1.125)
TSP sulfates (lag 1)	1.050 (0.976,1.129)	1.044 (0.971,1.123)	1.031 (0.959,1.109)

[a] Corresponds to Table 7 in Lippmann and colleagues 2000, p 18.

Table 2. Relative Risks (95% CI below) for Mortality from Circulatory Causes for PM Indices per 5th-to-95th Percentile Increment, 1985–1990[a]

	Original GAM (default)	GAM (stringent)	GLM
PM_{10} (lag 3)	1.023 (0.991,1.057)	1.010 (0.978,1.043)	1.005 (0.974,1.038)
TSP (lag 3)	1.014 (0.984,1.045)	1.004 (0.974,1.034)	0.999 (0.969,1.029)
TSP-PM_{10} (lag 1)	1.009 (0.977,1.043)	1.008 (0.976,1.041)	1.006 (0.974,1.039)
TSP sulfates (lag 2)	1.019 (0.991,1.048)	1.009 (0.982,1.038)	1.008 (0.980,1.036)

[a] Corresponds to Table 6 in Lippmann and colleagues 2000, p 18.

Table 4. Relative Risks (95% CI below) for Total Mortality for PM Indices per 5th-to-95th Percentile Increment, 1992–1994[a]

	Original GAM (default)	GAM (stringent)	GLM
$PM_{2.5}$ (lag 3)	1.045 (0.991,1.102)	1.027 (0.974,1.083)	1.029 (0.976,1.085)
$PM_{10-2.5}$ (lag 1)	1.038 (0.988,1.090)	1.031 (0.982,1.082)	1.027 (0.979,1.078)
PM_{10} (lag 1)	1.045 (0.989,1.102)	1.034 (0.980,1.091)	1.032 (0.978,1.089)
H^+ (lag 1)	1.010 (0.985,1.036)	1.007 (0.982,1.034)	1.006 (0.981,1.032)
$SO_4^=$ (lag 1)	1.024 (0.981,1.069)	1.015 (0.972,1.059)	1.012 (0.970,1.056)

[a] Corresponds to Table 13 in Lippmann and colleagues 2000, p 26.

Table 5. Relative Risks (95% CI below) for Mortality from Circulatory Causes for PM Indices per 5th-to-95th Percentile Increment, 1992–1994[a]

	Original GAM (default)	GAM (stringent)	GLM
$PM_{2.5}$ (lag 1)	1.046 (0.967,1.131)	1.032 (0.954,1.116)	1.029 (0.952,1.113)
$PM_{10-2.5}$ (lag 1)	1.075 (1.000,1.155)	1.064 (0.990,1.143)	1.058 (0.985,1.137)
PM_{10} (lag 1)	1.070 (0.987,1.160)	1.055 (0.973,1.142)	1.050 (0.969,1.137)
H^+ (lag 0)	1.015 (0.978,1.054)	1.014 (0.976,1.053)	1.012 (0.975,1.051)
$SO_4^=$ (lag 0)	1.018 (0.955,1.085)	1.013 (0.950,1.080)	1.009 (0.946,1.075)

[a] Corresponds to Table 13 in Lippmann and colleagues 2000, p 26.

Table 7. Relative Risks (95% CI below) for Hospital Admissions for Pneumonia Among the Elderly for PM Indices per 5th-to-95th Percentile Increment, 1992–1994[a]

	Original GAM (default)	GAM (stringent)	GLM
$PM_{2.5}$ (lag 1)	1.185 (1.053,1.332)	1.154 (1.027,1.298)	1.149 (1.022,1.292)
$PM_{10-2.5}$ (lag 1)	1.114 (1.006,1.233)	1.095 (0.990,1.211)	1.107 (1.00,1.226)
PM_{10} (lag 1)	1.219 (1.084,1.372)	1.185 (1.054,1.332)	1.190 (1.057,1.338)
H^+ (lag 3)	1.060 (1.005,1.118)	1.049 (0.994,1.107)	1.049 (0.994,1.107)
$SO_4^=$ (lag 1)	1.156 (1.050,1.273)	1.128 (1.025,1.242)	1.123 (1.020,1.235)

[a] Corresponds to Table 14 in Lippmann and colleagues 2000, p 26.

Table 6. Relative Risks (95% CI below) for Mortality from Respiratory Causes for PM Indices per 5th-to-95th Percentile Increment, 1992–1994[a]

	Original GAM (default)	GAM (stringent)	GLM
$PM_{2.5}$ (lag 0)	1.033 (0.855,1.248)	1.033 (0.854,1.249)	1.045 (0.864,1.265)
$PM_{10-2.5}$ (lag 2)	1.071 (0.913,1.257)	1.067 (0.909,1.253)	1.061 (0.904,1.245)
PM_{10} (lag 0)	1.080 (0.896,1.301)	1.077 (0.893,1.298)	1.081 (0.896,1.303)
H^+ (lag 1)	1.028 (0.938,1.128)	1.026 (0.936,1.126)	1.025 (0.934,1.124)
$SO_4^=$ (lag 3)	1.066 (0.908,1.251)	1.056 (0.899,1.240)	1.047 (0.892,1.230)

[a] Corresponds to Table 13 in Lippmann and colleagues 2000, p 26.

Table 8. Relative Risks (95% CI below) for Hospital Admissions for Chronic Obstructive Pulmonary Disease Among the Elderly for PM Indices per 5th-to-95th Percentile Increment, 1992–1994[a]

	Original GAM (default)	GAM (stringent)	GLM
$PM_{2.5}$ (lag 3)	1.080 (0.933,1.251)	1.043 (0.902,1.207)	1.004 (0.869,1.161)
$PM_{10-2.5}$ (lag 3)	1.089 (0.960,1.236)	1.083 (0.954,1.229)	1.103 (0.970,1.253)
PM_{10} (lag 3)	1.098 (0.946 1.274)	1.066 (0.920,1.235)	1.047 (0.904,1.213)
H^+ (lag 3)	1.067 (1.000,1.138)	1.055 (0.988,1.126)	1.039 (0.972,1.111)
$SO_4^=$ (lag 3)	1.060 (0.938,1.198)	1.032 (0.914,1.166)	0.990 (0.878,1.117)

[a] Corresponds to Table 14 in Lippmann and colleagues 2000, p 26.

Table 9. Relative Risks (95% CI below) for Hospital Admissions for Ischemic Heart Disease Among the Elderly for PM indices per 5th-to-95th Percentile Increment, 1992–1994[a]

	Original GAM (default)	GAM (stringent)	GLM
$PM_{2.5}$ (lag 2)	1.063 (0.980,1.153)	1.053 (0.971,1.143)	1.043 (0.961,1.131)
$PM_{10-2.5}$ (lag 2)	1.101 (1.026,1.181)	1.098 (1.023,1.178)	1.078 (1.004,1.157)
PM_{10} (lag 2)	1.091 (1.005,1.184)	1.082 (0.997,1.175)	1.063 (0.980,1.154)
H^+ (lag 2)	1.027 (0.991,1.065)	1.025 (0.989,1.063)	1.023 (0.987,1.061)
$SO_4^=$ (lag 2)	1.026 (0.961,1.095)	1.018 (0.954,1.087)	1.014 (0.949,1.082)

[a] Corresponds to Table 14 in Lippmann and colleagues 2000, p 27.

Table 11. Relative Risks (95% CI below) for Hosptial Admissions for Heart Failure Among the Elderly for PM Indices per 5th-to-95th Percentile Increment, 1992–1994[a]

	Original GAM (default)	GAM (stringent)	GLM
$PM_{2.5}$ (lag 1)	1.133 (1.034,1.241)	1.117 (1.020,1.224)	1.100 (1.004,1.205)
$PM_{10-2.5}$ (lag 0)	1.050 (0.968,1.138)	1.042 (0.962,1.130)	1.047 (0.966,1.135)
PM_{10} (lag 0)	1.099 (1.002,1.206)	1.094 (0.997,1.200)	1.086 (0.990,1.191)
H^+ (lag 0)	1.039 (0.992,1.088)	1.036 (0.989,1.085)	1.030 (0.983,1.079)
$SO_4^=$ (lag 0)	1.091 (1.012,1.176)	1.084 (1.006,1.168)	1.073 (0.996,1.156)

[a] Corresponds to Table 14 in Lippmann and colleagues 2000, p 27.

Table 10. Relative Risks (95% CI below) for Hospital Admissions for Dysrhythmias Among the Elderly for PM Indices 5th-to-95th Percentile Increment, 1992–1994[a]

	Original GAM (default)	GAM (stringent)	GLM
$PM_{2.5}$ (lag 1)	1.047 (0.907,1.208)	1.046 (0.906,1.207)	1.038 (0.900,1.198)
$PM_{10-2.5}$ (lag 0)	1.002 (0.882,1.138)	1.001 (0.881,1.138)	1.000 (0.880,1.137)
PM_{10} (lag 1)	1.030 (0.891,1.192)	1.029 (0.889,1.191)	1.020 (0.881,1.181)
H^+ (lag 0)	1.043 (0.974,1.117)	1.043 (0.974,1.117)	1.036 (0.968,1.110)
$SO_4^=$ (lag 1)	1.032 (0.921,1.157)	1.031 (0.920,1.155)	1.024 (0.913,1.148)

[a] Corresponds to Table 14 in Lippmann and colleagues 2000, p 27.

Table 12. Relative Risks (95% CI below) for Hospital Admissions for Stroke Among the Elderly for PM Indices per 5th-to-95th Percentile Increment, 1992–1994[a]

	Original GAM (default)	GAM (stringent)	GLM
$PM_{2.5}$ (lag 0)	1.026 (0.925,1.139)	1.028 (0.926,1.140)	1.014 (0.914,1.125)
$PM_{10-2.5}$ (lag 1)	1.047 (0.955,1.148)	1.048 (0.956,1.149)	1.054 (0.961,1.155)
PM_{10} (lag 1)	1.049 (0.944,1.165)	1.051 (0.946,1.167)	1.045 (0.941,1.161)
H^+ (lag 1)	1.024 (0.977,1.074)	1.025 (0.978,1.075)	1.024 (0.976,1.074)
$SO_4^=$ (lag 1)	1.015 (0.934,1.104)	1.018 (0.936,1.107)	1.013 (0.932,1.101)

[a] Corresponds to Table 14 in Lippmann and colleagues 2000, p 27.
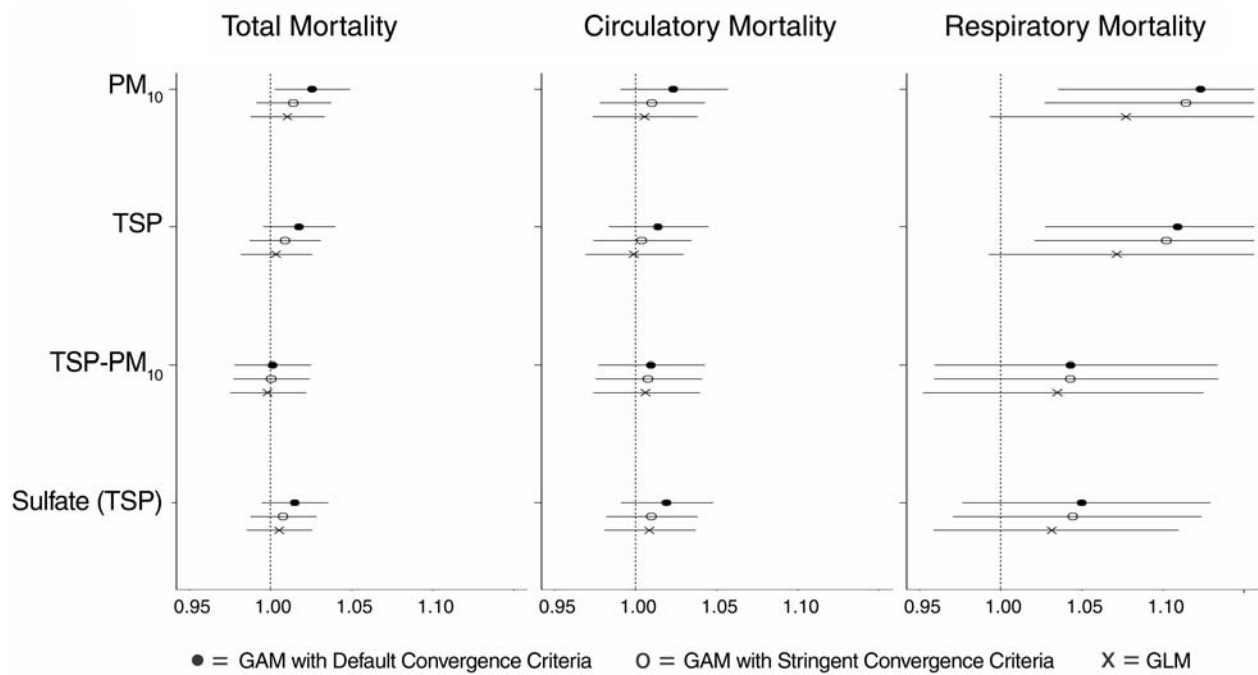
Figure 4. Relative risks for total mortality, mortality from circulatory causes, and mortality from respiratory causes per 5th-to-95th percentile increment in PM indices for 1985–1990 from GAM with default convergence criteria, GAM with stringent convergence criteria, and GLM with natural splines.
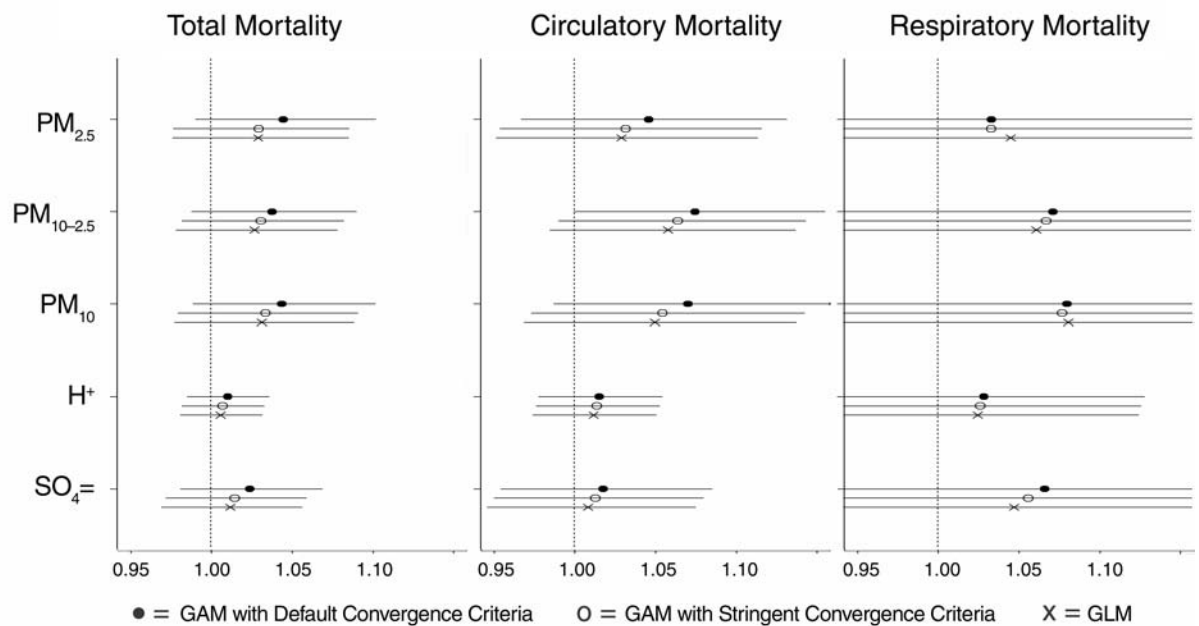
Figure 5. Relative risks for total mortality, mortality from circulatory causes, and mortality from respiratory causes per 5th-to-95th percentile increment in PM indices for 1992–1994 from GAM with default convergence criteria, GAM with stringent convergence criteria, and GLM with natural splines.
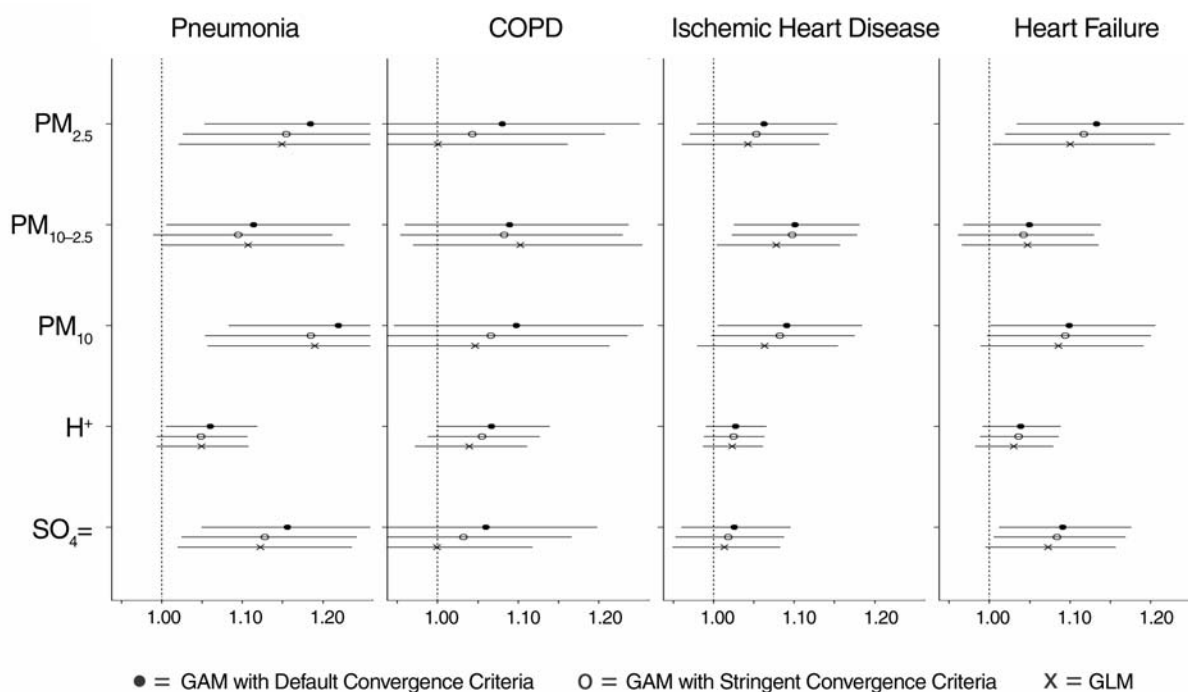
Figure 6. Hospital admissions among the elderly. Relative risks per 5th-to-95th percentile increment in PM indices for 1992–1994 from GAM with default convergence criteria, GAM with stringent convergence criteria, and GLM with natural splines. COPD = chronic obstructive pulmonary disease.

DISCUSSION

Reanalysis of the Detroit data using the original GAM with more stringent convergence criteria generally, but not always, resulted in reduced RR estimates. The extent of reduction appeared to be independent of the magnitude of RR estimates. The median difference in RRs per 5th-to-95th percentile increment of pollution index distribution was 0.0071. GLM analyses with natural splines that approximate the GAMs resulted in further but smaller reductions in RR estimates compared with those from the GAM with stringent convergence criteria. The median difference in RRs per 5th-to-95th percentile increment between the GAM with stringent convergence criteria and the GLM was 0.0028. These reductions in RR estimates did not differ across PM components. Therefore, the conclusions of the original study regarding the lack of relative role of PM by size and chemical characteristics remain the same.

Additional analyses using more common, simultaneous inclusion of all covariates resulted in generally larger but more variable RR estimates. Since the original two-step approach may have prevented the pollution variables from competing with other covariates, the approach used in the original analysis may give conservative pollution effect estimates. Additional sensitivity analyses using alternative weather model specifications that are similar to those used in the literature and with varying degrees of freedom for temporal trends resulted in PM_{10} coefficients over ranges that varied by a factor of two to three. These differences could be much larger than those observed in the revised analysis. Since we do not know exactly which weather model is correct or what extent of smoothing is appropriate, future analyses of this type of data may need to consider several sensitivity analyses using alternative weather models and varying extent of temporal smoothing.

Weather model specification and the extent of temporal smoothing are not the only factors that can change pollution RR estimates. Others may include the location of monitors, choice of lags, and considerations of distributed lags. These factors can cause differences that vary by up to a factor of two in estimated pollution coefficients. More comprehensive evaluations of relative impacts of these factors on pollution risk estimates are needed in the future.

RR per 5th-to-95th % PM$_{2.5}$ (black) and PM$_{10-2.5}$ (empty) Simultaneously Included in the Model

Circle: GAM with Default Convergence Criteria   Triangle: GAM with Stringent Convergence Criteria
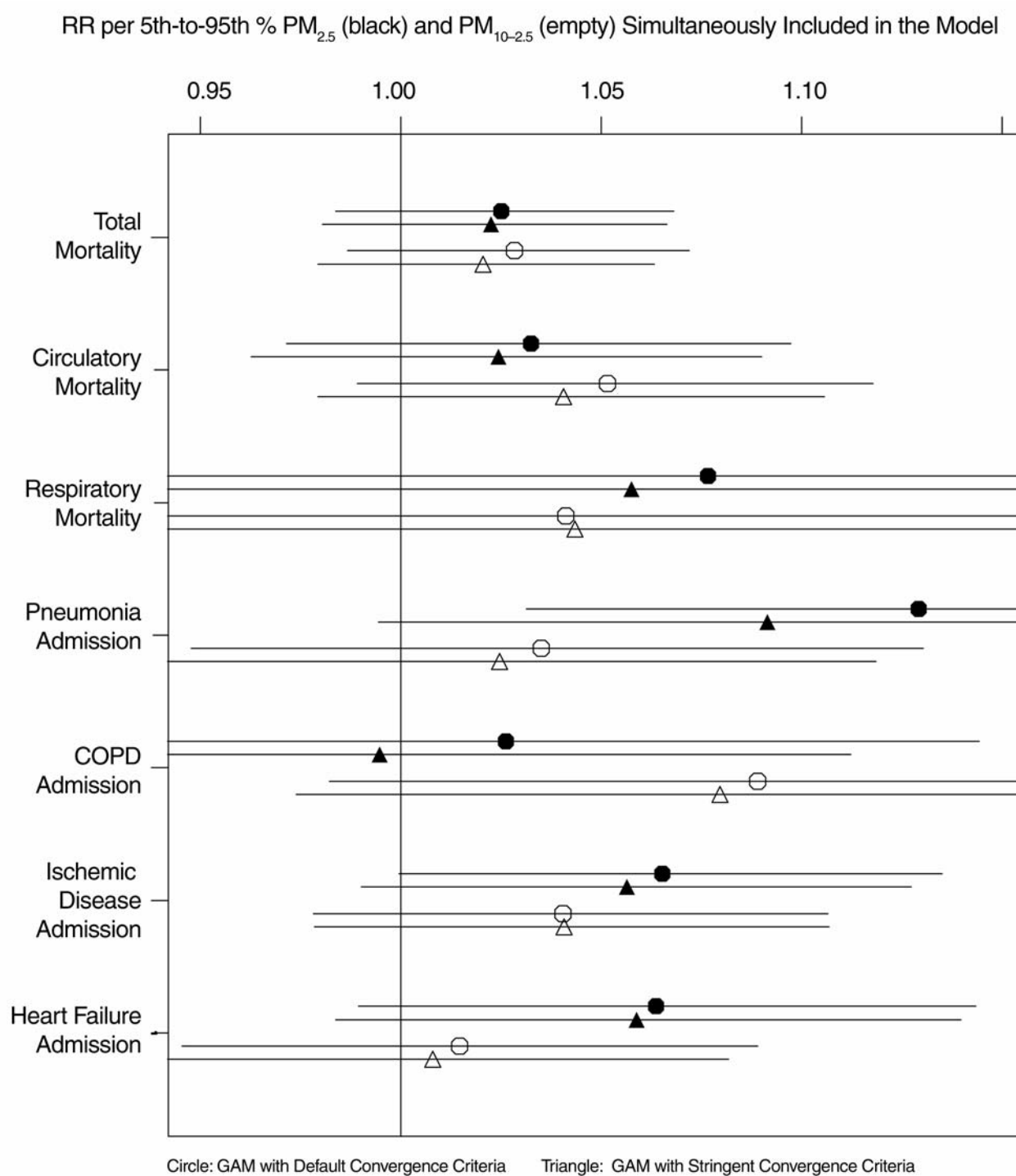
Figure 7. Mortality and elderly hospital admissions. Relative risks (RRs) per 5th-to-95th percentile increment in PM$_{2.5}$ and PM$_{10-2.5}$, included in the model simultaneously for 1992–1994, from GAM with default or stringent convergence criteria. COPD = chronic obstructive pulmonary disease.
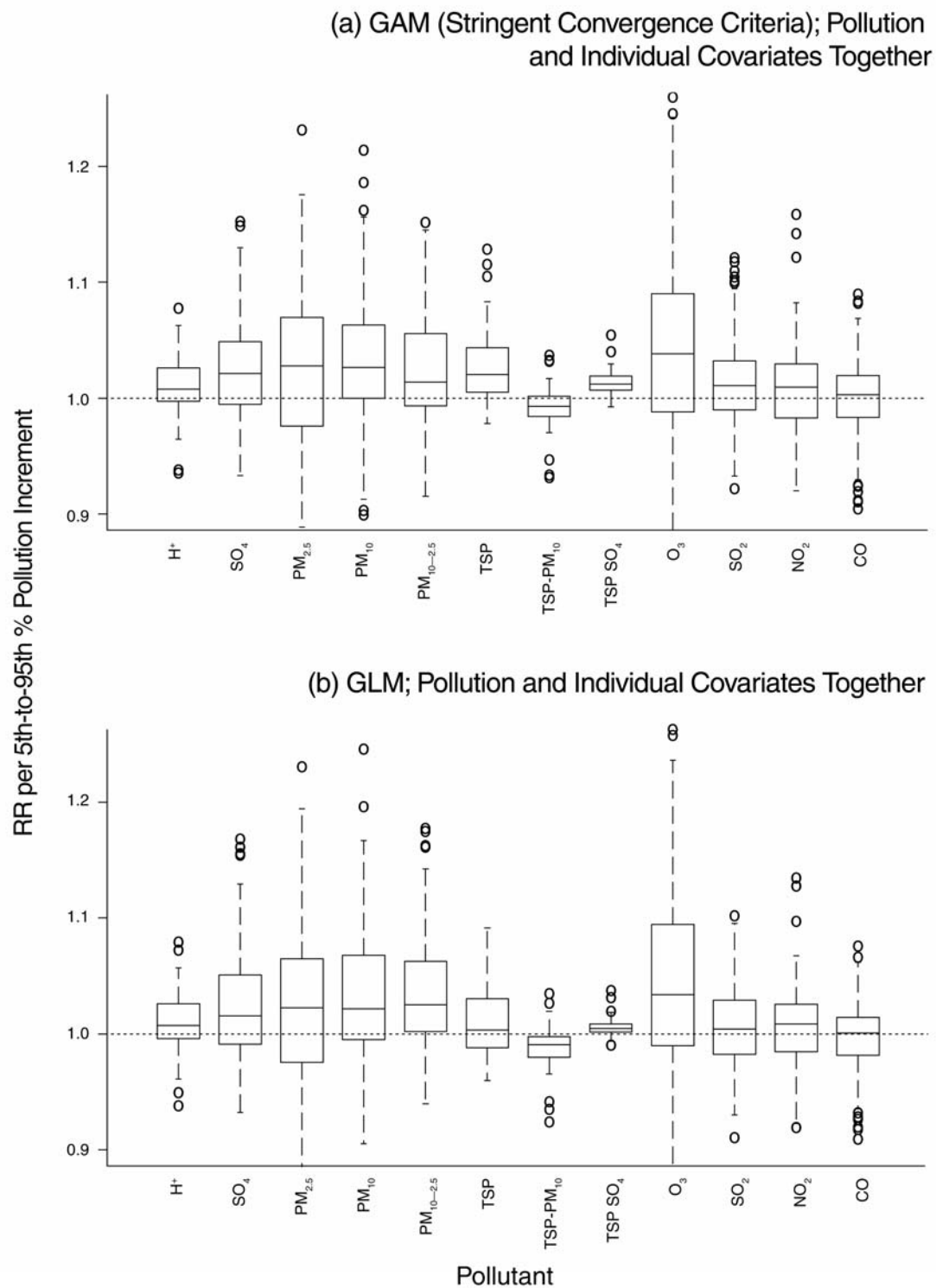
Figure 8. Box plots of all 1220 relative risks (RRs) for all health outcomes by pollutant for (a) GAM (stringent convergence criteria) with pollution and individual covariates together and (b) GLM with pollution and individual covariates together.
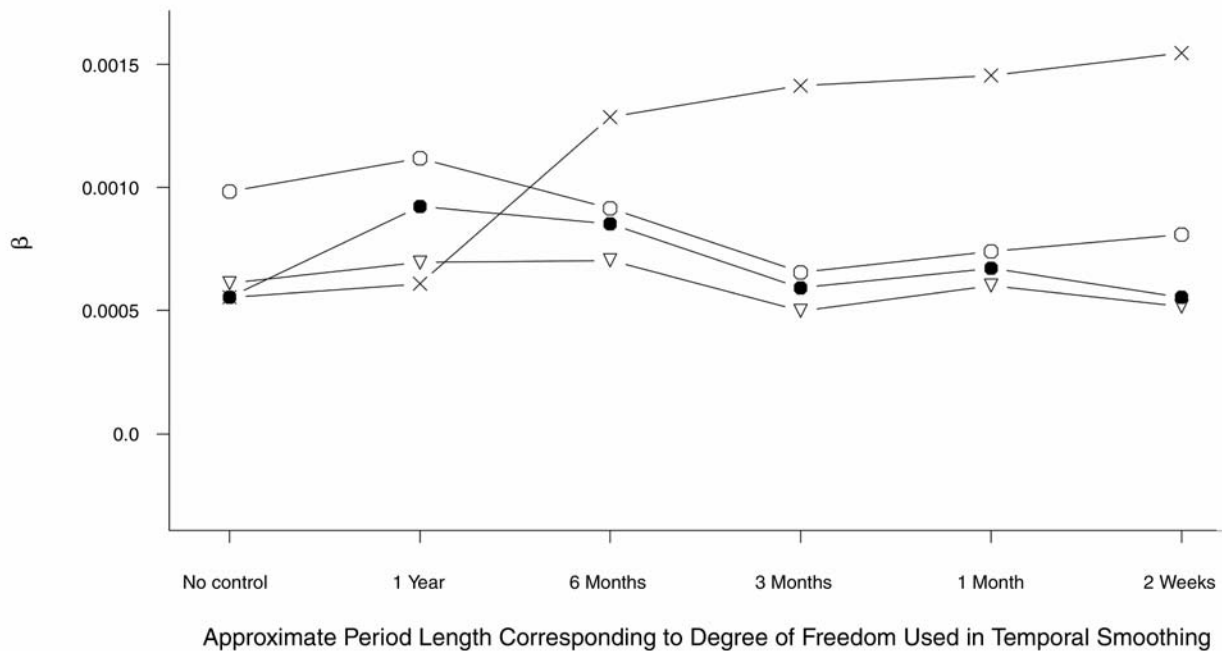
Figure 9. PM$_{10}$ (lag 1 day) coefficient (β) for total mortality, for 1992–1994, as a function of alternative weather models and varying degrees of freedom for fitting temporal trends using natural splines. White circle: natural splines of same-day temperature and same-day dewpoint, both with df = 2; black circle: natural splines of same-day temperature (df = 2), the average of temperatures lagged 1 through 3 days (df = 2), and hot-and-humid day indicator; white triangle: natural splines of the same-day temperature (df = 6), the average of temperatures lagged 1 through 3 days (df = 6), same-day dewpoint (df = 3), and the average of dewpoints lagged 1 through 3 days (df = 3); x: no adjustment for weather.
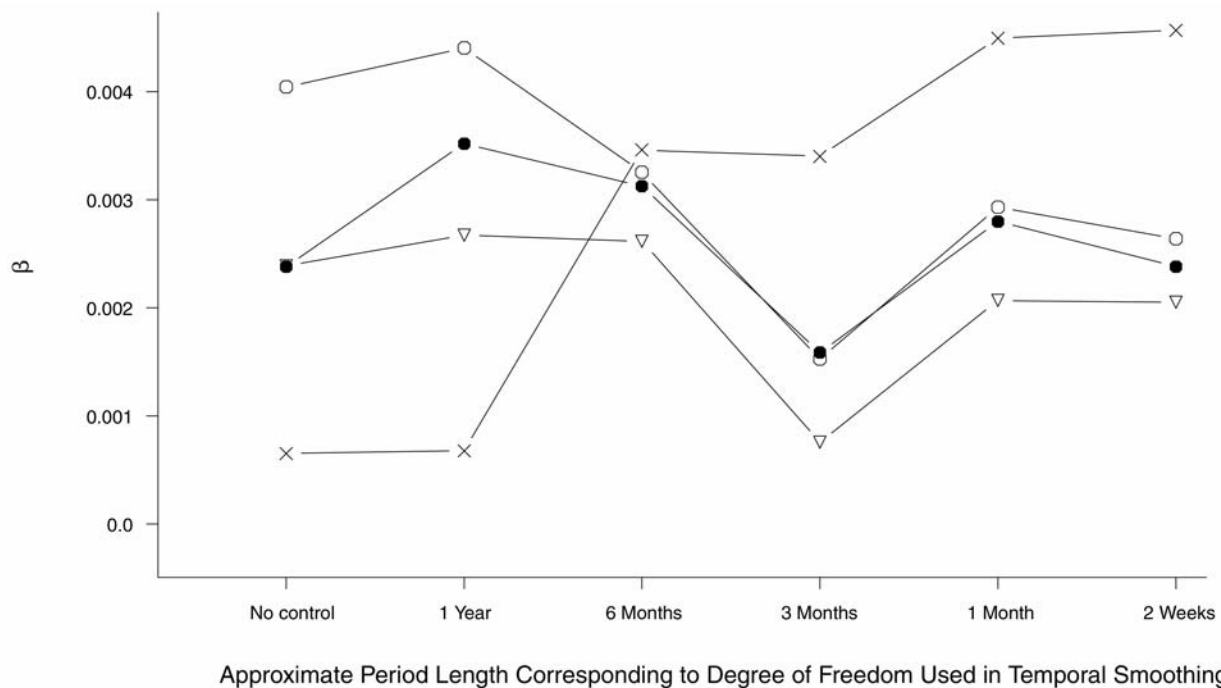
Figure 10. PM$_{10}$ (lag 1 day) coefficient (β) for hospital admissions for pneumonia among the elderly, for 1992–1994, as a function of alternative weather models and varying degrees of freedom for fitting temporal trends using natural splines. White circle: natural splines of same-day temperature and same-day dewpoint, both with df = 2; black circle: natural splines of same-day temperature (df = 2), the average of temperatures lagged 1 through 3 days (df = 2); white triangle: natural splines of the same-day temperature (df = 6), the average of temperatures lagged 1 through 3 days (df = 6), same-day dewpoint (df = 3), and the average of dewpoints lagged 1 through 3 days (df = 3); x: no adjustment for weather.

REFERENCES*

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On generalized additive models in time series studies of air pollution and health. Am J Epidemiol 156:1–11.

Lippmann M, Ito K, Nádas A, Burnett RT. 2000. Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations. Research Report 95. Health Effects Institute, Cambridge MA.

Samet JM, Dominici F, Zeger SL, Schwartz J, Dockery DW. 2000a. The National Morbidity, Mortality and Air Pollution Study, Part I: Methods and Methodologic Issues. Research Report 94. Health Effects Institute, Cambridge MA.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwartz J, Zanobetti A. 2000b. The National Morbidity, Mortality and Air Pollution Study, Part II: Mortality and Morbidity from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Schwartz J, Dockery DW, Neas LM. 1996. Is daily mortality associated specifically with fine particles? J Air Waste Manage Assoc 46:927–939.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
CI	confidence interval
CO	carbon monoxide
COPD	chronic obstructive pulmonary disease
df	degree of freedom
GAM	generalized additive model
GLM	generalized linear model
LOESS	locally weighted smoother
NO_2	nitrogen dioxide
O_3	ozone
PM	particulate matter
PM_{10}	PM less than 10 μm in diameter
$PM_{10-2.5}$	PM_{10} minus $PM_{2.5}$
$PM_{2.5}$	PM less than 2.5 μm in diameter
RR	relative risk
SO_2	sulfur dioxide
$SO_4{}^{2-}$	sulfate
TSP	total suspended particles

* Bold type identifies publication containing the original analyses revised in this short communication report.

Sensitivity Analysis of Various Models of Short-Term Effects of Ambient Particles on Total Mortality in 29 Cities in APHEA2

Klea Katsouyanni, Giota Touloumi, Evangelia Samoli, Yannis Petasakis, Antonis Analitis, Alain Le Tertre, Giuseppe Rossi, Denis Zmirou, Ferran Ballester, Azedine Boumghar, Hugh Ross Anderson, Bodgan Wojtyniak, Anna Paldy, Rony Braunstein, Juha Pekkanen, Christian Schindler, and Joel Schwartz

ABSTRACT

Under the APHEA2 (Air Pollution and Health: A European Approach *2) project we have reported results on the short-term effects of ambient particles on total nonaccidental mortality, using data from 29 European cities. The original method of analysis applied generalized additive models (GAM) with locally weighted smoothers (LOESS) for seasonality and meteorologic variables to individual city data and applied second-stage regression models to combine city-specific effect estimates and explore heterogeneity. After problems were recently identified for GAM for some types of analyses, we reanalyzed the original data using GAM with more stringent convergence criteria or using one of two parametric approaches: natural spline (ns) and penalized spline (ps).

We found that the original pooled effect estimate (b = 0.000617, SE = 0.000106) was reduced by 4% when more stringent GAM convergence criteria were applied (b = 0.000593, SE = 0.000103), by 33.6% when ns were applied (b = 0.000410, SE = 0.000091), and by 10.7% when ps were applied (b = 0.000550, SE = 0.000097); while the standard error estimates were practically unchanged. The effect modification patterns originally reported (higher PM_{10} [particulate matter less than 10 μm in diameter] effects in cities with higher nitrogen dioxide [NO_2] concentration, in warmer cities, in populations with higher age-standardized mortality rates, and in Southern and Northwestern cities compared to Central-eastern ones) were preserved when parametric methods for adjusting confounders were applied.

INTRODUCTION

Within the framework of the APHEA2 project we have reported results on the short-term effects of ambient particles on the total number of natural deaths, using data from 29 European cities (Katsouyanni et al 2001). For the analysis, a hierarchical modeling approach was adopted and implemented in two stages: individual city data were modeled and then these results were combined to estimate an overall effect or to explain heterogeneity and identify appropriate effect modifiers. Data for each city were analyzed separately, using generalized additive Poisson regression models with LOESS for seasonal patterns, temperature and humidity (Katsouyanni et al 2001).

Recently Dominici and colleagues (2002) identified a problem with the routine application of GAMs using S-Plus software with the default convergence criteria. In the National Morbidity, Mortality, and Air Pollution Study (NMMAPS), applying more stringent convergence criteria in S-Plus resulted in reducing by approximately 50% the mortality increase associated with an increase in PM_{10} of 10 μg/m^3. The same reduction was obtained when Poisson regression with parametric nonlinear adjustments (ns) for confounding factors was applied. Independently, Ramsay and colleagues (2003) found that the S-Plus GAM function underestimates the corresponding variances of the model parameters under certain conditions. This latter problem, however, does not seem to affect multicity studies or, more specifically, the pooled estimate (Daniels et al 2002).

In response to the above findings, we reanalyzed our data using the more stringent convergence criteria in S-Plus (Katsouyanni et al 2001). Responding to the request of EPA (US Environmental Protection Agency), we present that

---

reanalysis as well as a reanalysis of the data using Poisson regression with ns to adjust seasonal patterns and meteorologic variables. We also repeated the analysis using ps, an approach that is more flexible than ns but that at the same time maintains the advantages of a parametric method.

DATA AND METHODS

We used data from 29 European cities or areas: Athens (Greece); Barcelona, Bilbao, Madrid, and Valencia (Spain); Basel, Geneva, and Zurich (Switzerland); Birmingham and London (United Kingdom); Budapest (Hungary); Cracow, Lodz, Poznan, and Wroclaw (Poland); Dublin (Ireland); Erfurt (Germany); Helsinki (Finland); Ljubljana (Slovenia); Lyon, Marseille, and Paris (France); Milan, Rome, and Torino (Italy); Prague and Teplice (Czech Republic); Stockholm (Sweden); and Tel Aviv (Israel). Daily information was available for more than 5 years in each city for ambient particle concentrations (PM_{10} for 21 cities, black smoke for 14 cities) and levels of other pollutants and confounders. The outcome variable was the total nonaccidental daily number of deaths. Data were also collected for city (or area) characteristics describing geographic location, environmental and climatologic conditions, and the health status of the population. The role of these variables as potential effect modifiers was explored. More details on the data are provided in Katsouyanni and colleagues (2001).

Our original analysis approach included two modeling stages. First, the data for each city was analyzed separately, to allow for local differences, according to a predefined standardized method that resulted in a city-specific model. Briefly, GAM Poisson regression, using LOESS to control for seasonal patterns and meteorologic parameters, was applied. The S-Plus default convergence criteria were used. More details on the confounders and exposure variables included in the models can be found in Katsouyanni and colleagues (2001).

In the second stage of analysis, city-specific effect estimates were regressed on city-specific covariates to obtain an overall estimate and to explore heterogeneity across cities. We applied univariate (for single-pollutant city-specific models) and multivariate (for multipollutant city-specific models) regression models to investigate the role of potential effect modifiers. We estimated fixed-effects pooled regression coefficients and, in the presence of remaining significant heterogeneity, we also applied random-effects regression models to estimate the between-cities variance using the maximum likelihood method described by Berkey and colleagues (1995).

For the present reanalysis, we applied the first-stage models used in our original analyses but made the convergence criteria more stringent. More specifically, we set the maximum number of iterations to 1000 and the difference of two successive coefficients to 10^{-14}. We also applied two additional methods: using ns and ps parametric functions to adjust for seasonal patterns and meteorologic variables. As requested by the EPA, we used ns with the same number of degrees of freedom as those used in the original model (Table A.1). In the second method, proposed in this paper, we used ps with the same number of degrees of freedom as those in the original model. The latter approach is fully parametric but more flexible than ns or B-splines. The method is described in detail in Eilers and Marx (1996) and Marx and Eilers (1998).

RESULTS

Table 1 shows the individual-city and pooled β coefficients and standard errors for daily mortality associated with a 10-μg/m^3 increase in PM_{10} concentration, as estimated using the S-Plus GAM function with default convergence criteria (as in the original analysis) and more strict convergence criteria, ns, and ps. As reported before (Katsouyanni et al 2002), the differences between using the more stringent and the default convergence criteria were small: the pooled estimates were reduced by approximately 4% (Table 2); the change of individual city estimates ranged from −35.4% to +4.5%. The differences between ns and the original model were substantially larger. The fixed-effects pooled estimate was reduced by 38.5% and the random-effects pooled estimate by 33.6%. The changes in individual-city estimates ranged more widely, from −354.1% to +176.5% (Table 2). When ps were used to control for confounding effects, the estimated PM_{10} effects were closer to the original estimates. The pooled fixed-effects estimate and random-effects estimates were reduced by 16.0% and 10.7%, respectively. Similarly, the changes of individual-city effect estimates ranged from −185.8% to +36.6% (Table 2).

The standard errors of the coefficients from city-specific models generally increased when ns and ps were applied instead of the GAM method. However, the standard errors of the pooled estimates calculated with the random-effects model when ns and ps are applied were slightly lower than the corresponding standard errors from the GAMs.

Black smoke data were analyzed only with GAM with stricter convergence criteria or with ns parametric adjustments; ps was not applied. The results were similar to those for PM_{10} (Tables 3 and 4).

Table 1. PM_{10} (µg/m^3) Effects on Daily Total Number of Deaths: Results from GAM Using S-Plus Software Default Convergence Parameters (default), More Stringent Convergence Parameters (maximum number of iterations = 1000 and difference of two successive coefficients $< 10^{-14}$) (strict), Natural Splines[a] and Penalized Splines[a] for the APHEA2 Cities. All Models Adjusted for Seasonality, Long-Term Trends, Temperature, Humidity, Influenza, Day of the Week, Holidays, and Unusual Events if Necessary.

	Default		Strict		Natural Spline		Penalized Spline	
	β	SE	β	SE	β	SE	β	SE
City								
Athens	0.001534	0.000284	0.001473	0.000284	0.000914	0.000340	0.001311	0.000331
Barcelona	0.000928	0.000185	0.000856	0.000185	0.000274	0.000222	0.000575	0.000214
Basel	0.000412	0.000436	0.000407	0.000437	0.000288	0.000483	0.000462	0.000477
Birmingham	0.000282	0.000262	0.000289	0.000262	0.000233	0.000284	0.000305	0.000285
Budapest	0.000289	0.000462	0.000186	0.000461	−0.000734	0.000539	−0.000248	0.000520
Cracow	0.000135	0.000346	0.000118	0.000346	0.000245	0.000385	0.000155	0.000374
Erfurt	−0.000564	0.000394	−0.000570	0.000395	−0.000799	0.000431	−0.000465	0.000408
Geneva	−0.000103	0.000468	−0.000107	0.000469	−0.000284	0.000520	−0.000059	0.000512
Helsinki	0.000324	0.000427	0.000325	0.000427	0.000414	0.000446	0.000389	0.000442
London	0.000691	0.000175	0.000678	0.000175	0.000441	0.000197	0.000591	0.000197
Lyon	0.001353	0.000531	0.001356	0.000532	0.001362	0.000548	0.001554	0.000543
Madrid	0.000531	0.000238	0.000503	0.000237	0.000331	0.000270	0.000372	0.000269
Milano	0.001160	0.000189	0.001122	0.000188	0.000702	0.000239	0.000901	0.000228
Paris	0.000427	0.000230	0.000400	0.000229	0.000262	0.000273	0.000411	0.000277
Prague	0.000122	0.000183	0.000115	0.000183	0.000192	0.000198	0.000097	0.000201
Rome	0.001283	0.000270	0.001272	0.000270	0.001306	0.000302	0.001333	0.000301
Stockholm	0.000389	0.000863	0.000383	0.000863	0.000573	0.000887	0.000479	0.000883
Tel Aviv	0.000641	0.000259	0.000576	0.000258	0.000590	0.000282	0.000522	0.000276
Teplice	0.000641	0.000344	0.000639	0.000344	0.000758	0.000363	0.000876	0.000351
Torino	0.001046	0.000169	0.001006	0.000169	0.000568	0.000208	0.000938	0.000186
Zurich	0.000424	0.000370	0.000412	0.000370	0.000171	0.000399	0.000365	0.000388
Fixed Pooled Estimate								
	0.000682	0.000058	0.000654	0.000057	0.000420	0.000066	0.000573	0.000064
Random Pooled Estimate								
	0.000617	0.000106	0.000593	0.000103	0.000410	0.000091	0.000550	0.000097

[a] Using the same degree of freedom as were used in the original nonparametric city-specific model.

In order to evaluate the methods in terms of goodness of fit, we investigated city-specific and overall mean deviance for each method. All methods had very similar mean deviance.

We also explored the pattern of effect modification by the ns and ps methods. The estimated effects for different levels of the most important effect modifiers identified in the original analysis (long-term average NO_2, the ratio of PM_{10} to NO_2, long-term average temperature and humidity, age-standardized annual mortality rate, proportion of elderly, and area) (Katsouyanni et al 2002) preserved exactly the same pattern when the two parametric methods were used.

Table 2. Percent Difference in PM_{10} Regression Coefficients Using S-Plus GAM with Stringent Convergence Criteria (strict) or Using Parametric Adjustments with Natural Splines or Penalized Splines Compared with Values Obtained Using S-Plus GAM with Default Convergence Criteria

	% Difference in β Coefficients		
	Strict	Natural Spline	Penalized Spline
City			
Athens	−4.0	−40.4	−14.6
Barcelona	−7.8	−70.5	−38.1
Basel	−1.3	−30.1	12.2
Birmingham	2.4	−17.2	8.4
Budapest	−35.4	−354.1	−185.8
Cracow	−12.3	81.8	14.9
Erfurt	1.0	41.7	−17.6
Geneva	4.5	176.5	−42.8
Helsinki	0.5	27.9	20.1
London	−1.9	−36.3	−14.5
Lyon	0.2	0.7	14.9
Madrid	−5.3	−37.7	−30.0
Milan	−3.3	−39.5	−22.3
Paris	−6.2	−38.6	−3.8
Prague	−6.0	57.3	−20.2
Rome	−0.9	1.8	3.9
Stockholm	−1.5	47.2	23.2
Tel Aviv	−10.1	−7.9	−18.5
Tepliche	−0.3	18.3	36.6
Torino	−3.8	−45.7	−10.4
Zurich	−2.9	−59.7	−14.1
Fixed Pooled Estimate			
	−4.2	−38.5	−16.0
Random Pooled Estimate			
	−3.9	−33.6	−10.7

DISCUSSION

In this reanalysis, we found that applying GAMs with more stringent convergence criteria to the APHEA2 data does not practically affect the estimated effects of PM_{10} on mortality. The application of ns decreased the estimates by 30% to 40%, whereas the application of ps to control for confounding factors resulted in smaller decreases about 10% to 16%. Each method has advantages and problems, which we briefly discuss below.

GAMs have been extensively used in several research disciplines to model nonlinear relations (Schwartz et al 1996). In air pollution epidemiology, GAMs have replaced parametric methods as the standard approach. For example, sinusoidal terms for seasonal controls, used in APHEA1 (Katsouyanni et al 1997), were replaced by nonparametric smooth functions to control for season, long-term trends, and weather in APHEA2 (Samoli et al 2001).

The main reason for using GAM instead of parametric models is their increased flexibility in providing a more localized fit to the data. However, the calculation problems identified in the GAM functions (at least in S-Plus versions available before May 2002) may lead to biased estimates of regression coefficients and to underestimated associated standard errors. Regarding the latter issue, one multicity study has demonstrated that the standard error of the pooled estimate is not affected (Daniels et al 2002). In the present analysis, we showed that the standard error of the pooled estimate under the GAM random-effects model is similar (in fact, somewhat larger) to those estimated with parametric adjustment of confounders. However, this result does not hold for the standard errors of the pooled estimates under the fixed-effects model, which are underestimated by the GAM approach. Bootstraps in the city-specific models could be used to obtain unbiased standard errors.

Because of the problems identified in GAMs, Dominici and colleagues (2002) proposed the use of a fully parametric approach, using ns, to control for confounding effects. However, ns models may not be suitable. By definition the ns are restricted to be linear at the ends. This restriction can be a serious problem, particularly in modeling temperature in cities with extreme summer heat episodes. In addition, the results of ns models appear to be sensitive to the number and location of knots. In an initial exploratory analysis carried out in APHEA2, we found that the results change substantially when we move from an

Table 3. Black Smoke (µg/m^3) Effects on Daily Total Number of Deaths: Results from GAM Using S-Plus Software Default Convergence Parameters (default), More Stringent Convergence Parameters (maximum number of iterations = 1000 and difference of two successive coefficients < 10^{-14}) (strict), and Natural Splines[a] for the APHEA2 Cities. All Models Adjusted for Seasonality, Long-Term Trends, Temperature, Humidity, Influenza, Day of the Week, Holidays, and Unusual Events if Necessary

	Default		Strict		Natural Spline	
	β	SE	β	SE	β	SE
City						
Barcelona	0.001570	0.000273	0.001495	0.000273	0.001279	0.000315
Valencia	0.001342	0.000499	0.001294	0.000499	0.001329	0.000560
Marseille	0.001073	0.000361	0.001044	0.000361	0.000619	0.000418
Dublin	0.001038	0.000483	0.001041	0.000483	0.001221	0.000518
London	0.000929	0.000300	0.000904	0.000300	0.000423	0.000331
Bilbao	0.000813	0.000757	0.000735	0.000757	−0.000184	0.000836
Athens	0.000655	0.000117	0.000635	0.000117	0.000332	0.000133
Poznan	0.000624	0.000239	0.000625	0.000239	0.000504	0.000265
Paris	0.000383	0.000146	0.000365	0.000146	0.000280	0.000169
Birmingham	0.000342	0.000475	0.000338	0.000474	0.000222	0.000526
Wroclaw	0.000282	0.000228	0.000281	0.000228	0.000270	0.000244
Lodz	−0.000058	0.000211	−0.000057	0.000178	−0.000091	0.000194
Ljubljana	−0.000087	0.000609	−0.000100	0.000611	−0.000374	0.000659
Cracow	−0.000207	0.000212	−0.000214	0.000212	−0.000166	0.000232
Fixed Pooled Estimate						
	0.000512	0.000063	0.000476	0.000062	0.000307	0.000070
Random Pooled Estimate						
	0.000576	0.000133	0.000552	0.000131	0.000363	0.000112

[a] Using the same degree of freedom as were used in the original nonparametric city-specific model.

odd to an even number of knots. Apparently, this change is due to a change in the placement of knots: an odd number of knots will always result in placing one knot at the median of the distribution, whereas an even number of knots will not. The significance of this problem could be reduced if the number of knots was increased, but data overfitting would then be likely. The ps method, proposed here as an alternative, is also fully parametric and therefore maintains all the advantages of parametric methods. At the same time, it is more flexible than ns in providing a more localized fit to the data; thus it shares some advantages of nonparametric methods as well. The performance of the ps method in fitting real as well as simulated data under various scenarios has to be formally evaluated, but in our limited experience, the theoretical properties of ps are encouraging.

In the originally reported results (Katsouyanni et al 2001), we identified several effect modifiers that could partially explain the heterogeneity observed among city estimates. Significant heterogeneity was also present in the ns and ps reanalyses, as well as the same pattern of effects according to the levels of the identified effect modifiers.

In conclusion, our reanalyses suggest that the reported estimated effects in the APHEA2 project are reasonably robust to the application of alternative modeling strategies

Table 4. Percent Difference in Black Smoke Regression Coefficients Using S-Plus GAM with Stringent Convergence Criteria (Strict) or Using Parametric Adjustments with Natural Splines Compared with Values Obtained Using the S-Plus GAM with Default Convergence Criteria

	% Difference in β Coefficients	
	Strict	Natural Spline
City		
Barcelona	−4.8	−18.5
Valencia	−3.6	−1.0
Marseille	−2.7	−42.3
Dublin	0.3	17.7
London	−2.7	−54.5
Bilbao	−9.7	−122.6
Athens	−2.9	−49.3
Poznan	0.3	−19.2
Paris	−4.6	−26.8
Birmingham	−0.9	−35.1
Wroclaw	−0.3	−4.5
Lodz	−2.3	57.6
Ljubljana	14.9	332.0
Cracow	3.5	−19.7
Fixed Pooled Estimate		
	−7.0	−39.9
Random Pooled Estimate		
	−4.2	−37.1

REFERENCES*

Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. A random effects regression model for meta-analysis. Stat Med 1995; 14: 395–411.

Daniels M, Dominici F, Zeger S. Underestimation of standard errors in multi-site time-series studies of air pollution and health. Technical Report, 2002.

Dominici F, McDermott A, Zeger S, Samet J. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiol 2002; 156: 193–203

Eilers P and Marx BD. Flexible smoothing using b-splines and penalized likelihood. Statist Sci 1996; 11: 89–121.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiology 2001, 12: 521–31.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. Different convergence parameters applied to S-Plus GAM function. Epidemiology 2002; 13: 742–3.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak B, Sunyer J, Bacharova L, Schouten JP, Ponka A, Anderson HR. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. BMJ 1997; 314: 1658–1663.

Marx BD and Eilers P. Direct generalized additive modeling with penalized likelihood. Computational Statistics and Data Analysis 1998; 28: 193–209.

Ramsay T, Burnett R, Krewski D. The effect of concurvity in generalized additive models linking mortality and ambient air pollution. Epidemiology 2003, 14:18–23.

Samoli E, Schwartz J, Wojtyniak B, Touloumi G, Spix C, Balducci F, Medina S, Rossi G, Sunyer J, Bacharova L, Anderson HR, Katsouyanni K. Investigating regional differences in short-term effects of air pollution on daily mortality in the APHEA project: a sensitivity analysis. Env Health Perspect 2001;109:349–353.

Schwartz J, Spix C, Touloumi G, Bacharova L, Barumamdzadeh T, Le Tertre A, Piekarski T, Ponce de Leon A, Ponka A, Rossi G, Saez M, Schouten J. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. J Epidemiol Commun Health 1996; 50 (suppl 1): S3–S11.

* Bold type identifies publication containing the original analyses revised in this short communication report.

APPENDIX A

Table A.1. Variables and Degree of Freedom Included in City-Specific Models

City	Degree of Freedom						
	Days	Years	Trend	Temperature	Lag (temperature)	Humidity	Lag (humidity)
PM$_{10}$							
Athens	1805	5	15	4	6	linear	2
Barcelona	2190	7	23	4	3	linear	3
Basel	2190	6	11	4	linear	linear	linear
Birmingham	1640	5	15	4	linear	3	2
Budapest	1460	5	14	4	4	linear	—
Cracow	2555	7	15	6	linear	linear	2
Erfurt	1695	5	9	2	linear	linear	—
Geneva	2130	6	12	3	3	linear	—
Helsinki	1265	4	8	2	2	linear	—
London	1710	5	16	4	3	linear	2
Lyon	1790	5	12	4	linear	linear	2
Madrid	1460	4	16	4	6	linear	linear
Milano	2470	7	25	5	5	3	linear
Paris	1830	6	23	4	5	linear	linear
Prague	1325	5	15	3	2	2	linear
Rome	1685	5	8	9	6	2	linear
Stockholm	1005	7	5	2	linear	2	linear
Tel Aviv	1420	6	13	2	linear	linear	2
Teplice	1085	8	6	2	2	linear	linear
Torino	2530	7	21	2	4	2	linear
Zurich	2185	6	12	3	linear	linear	linear
Black Smoke							
Athens	1740	5	14	4	6	linear	2
Barcelona	2195	7	24	4	3	linear	3
Bilbao	1455	4	15	3	linear	linear	3
Birmingham	1825	5	22	4	linear	3	2
Cracow	2425	7	15	6	linear	linear	2
Dublin	2555	7	18	linear	2	linear	linear
Ljubljana	1825	5	10	2	4	linear	4
Lodz	2555	7	14	6	6	2	linear
London	1825	5	19	4	3	linear	2
Marseille	2190	6	13	4	linear	linear	linear
Paris	2185	6	23	4	6	linear	linear
Poznan	2485	7	17	5	6	3	linear
Valencia	1095	3	11	2	linear	linear	—
Wroclaw	2410	7	12	3	6	linear	—

ABBREVIATIONS AND OTHER TERMS

APHEA   Air Pollution and Health: A European Approach

EPA   US Environmental Protection Agency

GAM   generalized additive model

LOESS   locally weighted smoothers

NO_2 nitrogen dioxide

ns   natural spline

PM_{10} particulate matter less than 10 µm in diameter

ps   penalized spline

Replication of Reanalysis of Harvard Six-City Mortality Study

Rebecca J Klemm and Robert Mason

ABSTRACT

In their 1996 article, Schwartz, Dockery, and Neas reported that daily mortality was more strongly associated with concentrations of fine particulate matter ($PM_{2.5}$*) than with concentrations of larger particles (coarse mass [CM]) in six US cities (Schwartz et al 1996; hereafter this reference is termed *SDN*). Our revised analysis of their results was published in 2000 (Klemm et al 2000; hereafter this reference is termed *KMHND*). This short communication report compares the results presented in KMHND (and also SDN) with those rerun during 2002 with a more stringent convergence criterion for S-Plus generalized additive models (GAMs) with locally weighted smoothers (LOESS) and those run using S-Plus generalized linear models (GLMs) with natural spline smoothers for long-term trends and weather. Using the same data file created for KMHND, the new calculated estimated effects and the associated *t*-statistics of particulate air pollution on all-cause mortality as well as on mortality subgroups were, in general, reduced when using both GAM and GLM function methods. The figures illustrate changes in the estimated effects and their standard errors resulting from changes in the LOESS smoothing span and number of knots for natural splines.

INTRODUCTION

In their 1996 article, Schwartz, Dockery, and Neas reported that daily mortality was more strongly associated with concentrations of $PM_{2.5}$ than with concentrations of larger particles (CM) in six US cities (SDN). Because of the uniqueness of the concentration data used in these analyses

of mortality in six US cities and the importance of the SDN analyses in the US Environmental Protection Agency (EPA) standard-setting process, the Electric Power Research Institute undertook a reconstruction and reanalysis to replicate the original data and findings. Klemm Analysis Group, Inc, collaborated with two of the three SDN investigators to reconstruct and replicate the SDN analyses. The results were published in 2000 (KMHND). An independent oversight committee oversaw the process.

As part of the KMHND data reconstruction process, daily weather and daily counts of total and cause-specific deaths were reconstructed from original public records. Two-day average particulate series were constructed according to the SDN method (Table 2, KMHND). The reconstructed two-day average particulate air pollution and weather data were consistent with the summaries presented in SDN (Tables 5 and 6, KMHND; Tables 3 and 4, SDN). Daily counts of deaths in the reconstructed KMHND data file were lower than those in the SDN data file because of restrictions on residence and place of death. The KMHND reconstruction process also identified an administrative change in county codes that led to a higher number of daily deaths in St Louis.

The KMHND effects of particulate concentrations were estimated using the SDN GAM with LOESS for long-term trends and weather calculated in S-Plus.

The revised analyses involved rerunning the KMHND data file and model with a more stringent GAM convergence criterion as suggested by the EPA (L Grant, personal communication, 9/24/02). We also calculated the same model using a GLM function method with natural spline smoothers for long-term trends and weather for all of the mortality categories presented in KMHND. We compared the estimated effects calculated using GAM with LOESS with those calculated using GLM with natural spline smoothers as well as changes in the estimated effects and their standard errors resulting from changes in the LOESS smoothing span and number of knots for natural splines.

METHODS

Consistent with SDN and KMHND, we employed a Poisson regression model to estimate the association between mortality and particulate density, using S-Plus statistical software. Also consistent with the original analyses, all models included a single pollutant. The revised GAM used the same LOESS smoothing spans as published in KMHND and SDN: 50% for average daily temperature and average daily dewpoint temperature and 5% for time (study date). As in SDN and KMHND, six indicators for day of the week and a linear function for the particulate measurement completed the revised model.

The analyses presented here involved rerunning the KMHND data file and model using a more stringent GAM convergence criterion with the programming command suggested on page 3 of the EPA Communication (L Grant, personal communication, 9/24/02). We calculated the same model using GLM function methods with natural spline smoothers for the particulate mass concentrations and daily all-cause mortality and mortality subgroups that were presented in KMHND. The SDN-specified smoothing spans of 50% for average daily temperature and average daily dewpoint temperature each required approximately 3.5 degrees of freedom (*df*). To maintain consistency between the GAM and GLM function methods, we set the natural spline smoothers at 3 *df* (knots at the 25th and 75th percentile values) for average daily temperature and average dewpoint temperature and 38 *df* for time. (The *df* for time was one greater, 38, for GLM than for GAM LOESS, which was 37, to account for the GAM LOESS ½ *df* for both average daily temperature and average daily dewpoint.)

We also defined two GLMs that placed knots on a fixed-calendar basis using model degrees of freedom either fewer or greater than the GAM 51 *df* (6 for day of week, 1 for particulate concentration, 7 for weather, and 37 for time). The quarterly-knot GLM fixes one natural spline knot for time approximately every 90 days, at the 21st of March, June, September, and December; this GLM uses slightly fewer degrees of freedom for time. The monthly-knot GLM fixes one knot approximately every 30 days, at the 21st of each month; this GLM uses more degrees of freedom for time. The degree of freedom required for 4 knots per year and 12 knots per year varies by study area as a function of length of the data series.

Combined estimates were calculated as in KMHND. To create the combined estimate, each of the estimates specific to a study area was weighted by the inverse of the variance of the estimate.

RESULTS

Table 1 presents the estimated effects (and 95% confidence intervals [CIs] and *t*-statistics) on nonaccidental daily mortality associated with a 10 μg/m^3 increase in particulate mass concentration for $PM_{2.5}$, CM, PM_{15}/PM_{10} (PM < 15 or 10 μm in diameter), and sulfate (SO_4).

Table 1 shows that $PM_{2.5}$ estimated effects (combined and per study area) were lower or equal (1st decimal place) for GAM 2002 compared with KMHND-GAM (KMHND describes the differences between SDN and KMHND) and for GLM compared with GAM 2002. Boston showed the greatest differences in estimated positive effects between GLM and GAM. The *t*-statistics were lower for the GLMs. The results for CM, PM_{15}/PM_{10}, and SO_4 generally showed the same relative order of estimated effects as $PM_{2.5}$: SDN-GAM > KMHND-GAM > GAM 2002 > the three GLMs. The relative order of the GLM estimated effects varied by study area and particulate mass concentration.

Like in KMHND (page 1220), the combined estimates in Table 1 were dominated by St Louis (approximately 40–44%) and Boston (approximately 30–33%). The combined estimates for the GLMs resulted from greater weights for St Louis and Steubenville and decreased weight for Boston compared with the combined estimated for the GAMs.

Table 2 presents the results in the same format as Table 1 for mortality subgroups for $PM_{2.5}$: decedents at least 65 years of age (65+); all decedents due to chronic obstructive pulmonary disease (COPD); all decedents due to ischemic heart disease (IHD); and all decedents due to pneumonia. The overall pattern in the relative estimated effect size in Table 1 appears to be replicated in Table 2. (Estimated effect for CM, GAM 2002, exceeds that for CM, KMHND.)

Figures 1 through 4 graphically demonstrate the changes in estimated effects for $PM_{2.5}$ and their standard errors resulting from changes in the degree of freedom used to smooth time. Figure 1 (estimated effects) and Figure 2 (standard errors) relate to GAMs with the more stringent convergence criterion. Each curve consists of GAM estimated effects for a specific study area calculated using LOESS with spans for time of 0.5, 0.1, 0.05 (SDN model), 0.025, and 0.0125 (from left to right in Figures 1 and 2). Figures 3 and 4 are comparable figures for GLMs with calendar-specified fixed knots. Each curve consists of GLM estimated effects for a specific study area calculated using natural spline smoothers with knots every 90 days (4 knots/year), every 60 days, every 30 days (12 knots/year), every 20 days, and every 10 days (from left to right in Figures 3 and 4). In general, the estimated effects calculated using GAM were reduced as the degree of freedom to smooth time increased. This finding was not necessarily true of GLM.

Table 1. Estimated Changes in Nonaccidental Daily Mortality (95% CI) and *t*-statistics by Study Area and All Areas Combined Associated with a 10 μg/m³ Increase in Particulate Mass Concentration. Effect of Each Particulate Mass Measure Association was Estimated Separately with Control for Long-Term Trends and Weather.[a]

Source	df for Time[b]	PM$_{2.5}$	CM	PM$_{15}$ or PM$_{10}$	SO$_4$
Boston					
SDN-GAM		2.2% (1.5%, 2.9%) *t*=6.31	0.2% (−0.6%, 1.2%) *t*=0.58	1.2% (0.7%, 1.7%) *t*=4.86	Not included
KMHND-GAM	36.6	2.1% (1.4%, 2.8%) *t*=5.97	0.5% (−0.5%, 1.4%) *t*=0.99	1.2% (0.7%, 1.7%) *t*=4.84	3.8% (2.3%, 5.2%) *t*=5.15
GAM 2002	36.6	2.0% (1.3%, 2.7%) *t*=5.72	0.5% (−0.5%, 1.4%) *t*=0.96	1.2% (0.7%, 1.7%) *t*=4.65	3.5% (2.1%, 4.9%) *t*=4.92
GLM *df* mtch	38	1.6% (0.8%, 2.4%) *t*=3.84	0.2% (−1.0%, 1.3%) *t*=0.29	0.9% (0.3%, 1.5%) *t*=2.94	2.8% (1.0%, 4.5%) *t*=3.11
GLM 4kn/yr	28	1.6% (0.8%, 2.5%) *t*=3.91	0.7% (−0.4%, 1.8%) *t*=1.19	1.1% (0.5%, 1.6%) *t*=3.47	2.6% (0.8%, 4.3%) *t*=2.88
GLM 12kn/yr	81	0.8% (−0.1%, 1.6%) *t*=1.83	0.0% (−1.1%, 1.2%) *t*=0.02	0.4% (−0.2%, 1.0%) *t*=1.32	1.6% (−0.2%, 3.4%) *t*=1.74
Knoxville					
SDN-GAM		1.4% (0.2%, 2.6%) *t*=2.26	1.0% (−0.6%, 2.6%) *t*=1.20	0.9% (0.1%, 1.8%) *t*=2.21	Not included
KMHND-GAM	36.1	1.9% (0.6%, 3.3%) *t*=2.79	1.6% (−0.2%, 3.4%) *t*=1.77	1.4% (0.4%, 2.3%) *t*=2.88	0.6% (−2.0%, 3.3%) *t*=0.45
GAM 2002	36.1	1.7% (0.4%, 3.1%) *t*=2.53	1.4% (−0.4%, 3.2%) *t*=1.56	1.2% (0.3%, 2.1%) *t*=2.59	0.6% (−2.1%, 3.2%) *t*=0.42
GLM *df* mtch	38	1.3% (−0.2%, 2.9%) *t*=1.67	1.1% (−0.9%, 3.1%) *t*=1.10	1.0% (−0.1%, 2.0%) *t*=1.74	0.4% (−2.7%, 3.5%) *t*=0.24
GLM 4kn/yr	33	1.5% (−0.1%, 3.0%) *t*=1.85	1.2% (−0.8%, 3.2%) *t*=1.18	1.0% (−0.0%, 2.1%) *t*=1.91	0.7% (−2.4%, 3.7%) *t*=0.42
GLM 12kn/yr	96	1.1% (−0.5%, 2.7%) *t*=1.30	0.8% (−1.3%, 2.9%) *t*=0.73	0.7% (−0.4%, 1.9%) *t*=1.30	0.1% (−3.1%, 3.2%) *t*=0.03
St Louis					
SDN-GAM		1.1% (0.4%, 1.7%) *t*=3.17	0.2% (−0.7%, 1.1%) *t*=0.45	0.6% (0.1%, 1.0%) *t*=2.42	Not included
KMHND-GAM	36.4	0.8% (0.2%, 1.4%) *t*=2.62	0.1% (−0.8%, 1.0%) *t*=0.20	0.4% (−0.0%, 0.9%) *t*=1.91	0.6% (−0.6%, 1.7%) *t*=0.98
GAM 2002	36.4	0.8% (0.2%, 1.4%) *t*=2.54	0.0% (−0.9%, 0.9%) *t*=0.06	0.4% (−0.0%, 0.8%) *t*=1.79	0.6% (−0.6%, 1.7%) *t*=0.99
GLM *df* mtch	38	0.6% (−0.1%, 1.3%) *t*=1.61	−0.0% (−1.0%, 1.0%) *t*=−0.01	0.3% (−0.2%, 0.8%) *t*=1.06	0.5% (−0.8%, 1.8%) *t*=0.73
GLM 4kn/yr	30	0.5% (−0.2%, 1.2%) *t*=1.48	−0.2% (−1.2%, 0.8%) *t*=−0.45	0.2% (−0.3%, 0.7%) *t*=0.75	0.5% (−0.8%, 1.8%) *t*=0.73
GLM 12kn/yr	88	0.1% (−0.6%, 0.8%) *t*=0.26	−0.2% (−1.3%, 0.8%) *t*=−0.47	−0.0% (−0.5%, 0.5%) *t*=−0.13	−0.1% (−1.5%, 1.2%) *t*=−0.22

[a] SDN = Schwartz et al 1996; KMHND = Klemm et al 2000; CM = coarse mass.
SDN-GAM = GAM with LOESS for long-term trends and weather as published in SDN Table 5. SDN investigators provided previously unreported SO$_4$ results for the KMHND analysis.
KMHND-GAM = GAM with LOESS for long-term trends and weather as published in KMHND Table 7.
GAM 2002 = KMHDN-GAM with LOESS for long-term trends and weather but with the more stringent convergence criterion suggested by EPA
(S-Plus string epsilon=10e-8, maxit=1000, bf.epsilon=10e-8, bf.maxit=1000, trace=T).
GLM *df* mtch = GLM with natural splines for weather and long-term trends using 38 *df* for time to match the approximate GAM 2002 51 *df* for each study area.
GLM 4kn/yr = GLM with natural splines for weather and long-term trends using a knot placed on the 21st of March, June, September and December (approximately every 90 days).
GLM 12kn/yr = GLM with natural splines for weather and long-term trends using a knot placed on the 21st of each month (approximately every 30 days).
[b] For PM$_{2.5}$, the degree of freedom for time varied slightly for other particulate mass concentrations.

167

Table 1 (Continued). Estimated Changes in Nonaccidental Daily Mortality (95% CI) and *t*-statistics by Study Area and All Areas Combined Associated with a 10 μg/m³ Increase in Particulate Mass Concentration. Effect of Each Particulate Mass Measure Association was Estimated Separately with Control for Long-Term Trends and Weather.[a]

Source	*df* for Time[b]	PM₂.₅	CM	PM₁₅ or PM₁₀	SO₄
Steubenville					
SDN-GAM		1.0% (−0.1%, 2.1%) *t*=1.79	2.4% (0.5%, 4.3%) *t*=2.43	0.9% (0.1%, 1.6%) *t*=2.17	Not included
KMHND-GAM	36.8	0.6% (−0.6%, 1.9%) *t*=1.00	1.9% (−0.2%, 4.0%) *t*=1.77	0.6% (−0.3%, 1.4%) *t*=1.36	1.4% (−1.2%, 4.0%) *t*=1.06
GAM 2002	36.8	0.6% (−0.7%, 1.8%) *t*=0.90	1.8% (−0.3%, 3.9%) *t*=1.70	0.5% (−0.3%, 1.4%) *t*=1.26	1.3% (−1.3%, 3.9%) *t*=0.98
GLM *df* mtch	38	0.2% (−1.1%, 1.6%) *t*=0.37	1.4% (−0.8%, 3.6%) *t*=1.23	0.3% (−0.6%, 1.2%) *t*=0.72	0.7% (−2.0%, 3.5%) *t*=0.51
GLM 4kn/yr	35	0.2% (−1.1%, 1.5%) *t*=0.31	1.6% (−0.6%, 3.9%) *t*=1.42	0.3% (−0.6%, 0.7%) *t*=0.75	0.6% (−2.1%, 3.4%) *t*=0.45
GLM 12kn/yr	103	0.6% (−0.8%, 2.0%) *t*=0.81	1.5% (−0.9%, 3.9%) *t*=1.25	0.5% (−0.5%, 1.5%) *t*=1.03	1.2% (−1.7%, 4.1%) *t*=0.80
Portage					
SDN-GAM		1.2% (−0.3%, 2.8%) *t*=1.64	0.5% (−1.2%, 2.3%) *t*=0.57	0.7% (−0.4%, 1.7%) *t*=1.22	Not included
KMHND-GAM	36.9	0.9% (−0.8%, 2.5%) *t*=1.00	0.1% (−1.8%, 2.1%) *t*=0.11	0.4% (−0.7%, 1.5%) *t*=0.68	1.9% (−0.9%, 4.9%) *t*=1.31
GAM 2002	36.9	0.6% (−1.1%, 2.3%) *t*=0.72	−0.0% (−2.0%, 1.9%) *t*=−0.05	0.2% (−0.9%, 1.4%) *t*=0.37	1.6% (−1.3%, 4.4%) *t*=1.08
GLM *df* mtch	38	−0.7% (−2.6%, 1.2%) *t*=−0.71	−0.6% (−2.8%, 1.6%) *t*=−0.56	−0.5% (−1.8%, 0.8%) *t*=−0.77	−0.2% (−3.5%, 3.0%) *t*=−0.13
GLM 4kn/yr	36	−0.5% (−2.3%, 1.4%) *t*=−0.47	−0.4% (−2.5%, 1.8%) *t*=−0.35	−0.3% (−1.6%, 1.0%) *t*=−0.50	0.0% (−3.2%, 3.2%) *t*=0.01
GLM 12kn/yr	106	−0.9% (−2.9%, 1.1%) *t*=−0.89	−1.2% (−3.5%, 1.1%) *t*=−1.05	−0.9% (−2.2%, 5.1%) *t*=−1.23	−0.7% (−4.2%, 2.7%) *t*=−0.43
Topeka					
SDN-GAM		0.8% (−2.0%, 3.6%) *t*=0.53	−1.3% (−3.3%, 0.6%) *t*=−1.32	−0.5% (−2.0%, 0.9%) *t*=−0.67	Not included
KMHND-GAM	37.3	0.6% (−2.6%, 4.0%) *t*=0.38	−1.5% (−3.8%, 0.8%) *t*=−1.28	−0.7% (−2.3%, 1.1%) *t*=−0.75	−0.8% (−6.4%, 5.1%) *t*=−0.27
GAM 2002	37.3	0.6% (−2.7%, 3.9%) *t*=0.35	−1.5% (−3.8%, 0.8%) *t*=−1.26	−0.7% (−2.4%, 1.1%) *t*=−0.75	−0.8% (−6.6%, 5.0%) *t*=−0.27
GLM *df* mtch	38	−0.1% (−4.0%, 3.7%) *t*=−0.07	−1.7% (−4.4%, 0.9%) *t*=−1.30	−1.0% (−2.9%, 1.0%) *t*=−1.00	−1.5% (−8.0%, 5.1%) *t*=−0.44
GLM 4kn/yr	37	−0.2% (−4.0%, 3.6%) *t*=−0.09	−1.9% (−4.5%, 0.7%) *t*=−1.43	−1.1% (−3.0%, 0.9%) *t*=−1.10	−1.6% (−8.1%, 5.0%) *t*=−0.46
GLM 12kn/yr	108	−0.7% (−4.8%, 3.3%) *t*=−0.36	−3.0% (−5.9%, −0.0%) *t*=−1.98	−1.8% (−3.9%, 0.4%) *t*=−1.62	−0.6% (−7.8%, 6.5%) *t*=−0.17
Combined					
SDN-GAM		1.5% (1.1%, 1.9%) *t*=7.41	0.4% (−0.1%, 1.0%) *t*=1.48	0.8% (0.5%, 1.1%) *t*=5.84	2.2% (1.3%, 3.1%) *t*=4.67
KMHND-GAM		1.3% (0.9%, 1.7%) *t*=6.53	0.4% (−0.2%, 0.9%) *t*=1.43	0.8% (0.5%, 1.0%) *t*=5.27	1.6% (0.9%, 2.4%) *t*=4.20
GAM 2002		1.2% (0.8%, 1.6%) *t*=6.16	0.3% (−0.2%, 0.9%) *t*=1.20	0.7% (0.4%, 1.0%) *t*=4.89	1.5% (0.8%, 2.3%) *t*=3.98
GLM *df* mtch		0.8% (0.4%, 1.3%) *t*=3.61	0.1% (−0.5%, 0.7%) *t*=0.37	0.4% (0.1%, 0.8%) *t*=2.68	1.0% (0.1%, 1.9%) *t*=2.21
GLM 4kn/yr		0.8% (0.4%, 1.3%) *t*=3.63	0.2% (−0.4%, 0.8%) *t*=0.70	0.5% (0.1%, 0.8%) *t*=2.87	1.0% (0.1%, 1.9%) *t*=2.16
GLM 12kn/yr		0.4% (−0.1%, 0.8%) *t*=1.57	−0.2% (−0.8%, 0.5%) *t*=−0.46	0.1% (−0.2%, 0.5%) *t*=0.82	0.4% (−0.5%, 1.3%) *t*=0.85

[a] SDN = Schwartz et al 1996; KMHND = Klemm et al 2000; CM = coarse mass.
SDN-GAM = GAM with LOESS for long-term trends and weather as published in SDN Table 5. SDN investigators provided previously unreported SO₄ results for the KMHND analysis.
KMHND-GAM = GAM with LOESS for long-term trends and weather as published in KMHND Table 7.
GAM 2002 = KMHDN-GAM with LOESS for long-term trends and weather but with the more stringent convergence criterion suggested by EPA
(S-Plus string epsilon=10e-8,maxit=1000,bf.epsilon=10e-8,bf.maxit=1000,trace=T).
GLM *df* mtch = GLM with natural splines for weather and long-term trends using 38 *df* for time to match the approximate GAM 2002 51 *df* for each study area.
GLM 4kn/yr = GLM with natural splines for weather and long-term trends using a knot placed on the 21st of March, June, September and December (approximately every 90 days).
GLM 12kn/yr = GLM with natural splines for weather and long-term trends using a knot placed on the 21st of each month (approximately every 30 days).

[b] For PM₂.₅, the degree of freedom for time varied slightly for other particulate mass concentrations.

Table 2. Estimated Changes in Daily Mortality for Specific Mortality Categories (95% CI) and *t*-statistics by Study Area and All Areas Combined Associated with a 10 μg/m³ Increase in Fine Particulate Mass Concentration. Effect of Fine Association was Estimated Separately for Each Mortality Category with Control for Long-Term Trends and Weather[a]

Source	*df* for Time[b]	Age 65+, PM$_{2.5}$	COPD, PM$_{2.5}$	IHD, PM$_{2.5}$	Pneumonia, PM$_{2.5}$
Boston					
SDN-GAM		Not included	Not included	Not included	Not included
KMHND-GAM		Not included	Not included	Not included	Not included
GAM 2002	36.6	2.4% (1.6%, 3.1%) *t*=6.00	2.8% (−1.3%, 6.8%) *t*=1.34	2.7% (1.5%, 3.9%) *t*=4.41	5.9% (2.6%, 9.1%) *t*=3.56
GLM *df* mtch	38	1.9% (1.0%, 2.8%) *t*=4.11	2.0% (−2.7%, 6.7%) *t*=0.85	2.2% (0.8%, 3.7%) *t*=3.10	4.0% (0.2%, 7.8%) *t*=2.09
GLM 4kn/yr	28	1.9% (1.0%, 2.8%) *t*=4.11	2.2% (−2.5%, 6.9%) *t*=0.92	2.4% (1.0%, 3.9%) *t*=3.37	4.5% (0.8%, 8.3%) *t*=2.36
GLM 12kn/yr	81	1.1% (1.3%, 2.0%) *t*=2.23	1.6% (−3.5%, 6.7%) *t*=0.62	1.4% (−0.1%, 2.9%) *t*=1.77	1.1% (−2.9%, 5.1%) *t*=0.53
Knoxville					
SDN-GAM		Not included	Not included	Not included	Not included
KMHND-GAM		Not included	Not included	Not included	Not included
GAM 2002	36.1	1.8% (0.3%, 3.4%) *t*=2.30	0.7% (−6.5%, 7.8%) *t*=0.18	2.7% (0.1%, 5.2%) *t*=2.03	2.5% (−6.0%, 10.9%) *t*=0.59
GLM *df* mtch	38	1.4% (−0.4%, 3.1%) *t*=1.50	−1.2% (−9.7%, 7.2%) *t*=−0.28	2.3% (−0.7%, 5.2%) *t*=1.50	1.4% (−8.6%, 11.5%) *t*=0.28
GLM 4kn/yr	33	1.6% (−0.2%, 3.3%) *t*=1.73	−0.8% (−9.2%, 7.6%) *t*=−0.18	2.1% (−0.9%, 5.0%) *t*=1.38	0.9% (−9.0%, 10.8%) *t*=0.17
GLM 12kn/yr	96	1.0% (−0.8%, 2.8%) *t*=1.06	0.4% (−8.6%, 9.4%) *t*=0.09	2.2% (−0.9%, 5.3%) *t*=1.39	1.1% (−9.4%, 11.5%) *t*=0.20
St Louis					
SDN-GAM		Not included	Not included	Not included	Not included
KMHND-GAM		Not included	Not included	Not included	Not included
GAM 2002	36.4	0.7% (−0.0%, 1.4%) *t*=1.94	0.6% (−2.9%, 4.2%) *t*=0.35	1.3% (0.3%, 2.4%) *t*=2.45	1.1% (−2.5%, 4.7%) *t*=0.60
GLM *df* mtch	38	0.4% (−0.4%, 1.2%) *t*=0.99	0.3% (−3.5%, 4.1%) *t*=0.15	1.2% (0.0%, 2.4%) *t*=2.01	−0.0% (−4.1%, 4.0%) *t*=−0.02
GLM 4kn/yr	30	0.4% (−0.4%, 1.2%) *t*=0.97	0.3% (−3.5%, 4.0%) *t*=0.15	1.0% (−0.1%, 2.2%) *t*=1.75	0.7% (−3.4%, 4.7%) *t*=0.32
GLM 12kn/yr	88	−0.1% (−0.9%, 0.7%) *t*=−0.28	0.2% (−3.8%, 4.1%) *t*=0.09	0.4% (−0.8%, 1.7%) *t*=0.72	−0.8% (−5.0%, 3.4%) *t*=−0.38

Table continues next page

[a] SDN = Schwartz et al 1996; KMHND = Klemm et al 2000; COPD = chronic obstructive pulmonary disease; IHD = ischemic heart disease.
SDN-GAM = GAM with LOESS for long-term trends and weather as published in SDN Table 5. SDN investigators provided previously unreported SO$_4$ results for the KMHND analysis.
KMHND-GAM = GAM with LOESS for long-term trends and weather as published in KMHND Table 7.
GAM 2002 = KMHDN-GAM with LOESS for long-term trends and weather but with the more stringent convergence criterion suggested by EPA
  (S-Plus string epsilon=10e-8,maxit=1000,bf.epsilon=10e-8,bf.maxit=1000,trace=T).
GLM *df* mtch = GLM with natural splines for weather and long-term trends using 38 *df* for time to match the approximate GAM 2002 51 *df*
  for each study area.
GLM 4kn/yr = GLM with natural splines for weather and long-term trends using a knot placed on the 21st of March, June, September and December
  (approximately every 90 days).
GLM 12kn/yr = GLM with natural splines for weather and long-term trends using a knot placed on the 21st of each month (approximately every 30 days).

[b] For PM$_{2.5}$; the degree of freedom for time varied slightly for other particulate mass concentrations.

Table 2 (Continued). Estimated Changes in Daily Mortality for Specific Mortality Categories (95% CI) and *t*-statistics by Study Area and All Areas Combined Associated with a 10 μg/m³ Increase in Fine Particulate Mass Concentration. Effect of Fine Association was Estimated Separately for Each Mortality Category with Control for Long-Term Trends and Weather[a]

Source	*df* for Time[b]	Age 65+, $PM_{2.5}$	COPD, $PM_{2.5}$	IHD, $PM_{2.5}$	Pneumonia, $PM_{2.5}$
Steubenville					
SDN-GAM		Not included	Not included	Not included	Not included
KMHND-GAM		Not included	Not included	Not included	Not included
GAM 2002	36.8	1.5% (0.1%, 2.8%) *t*=2.10	8.1% (2.3%, 13.9%) *t*=2.72	0.8% (−1.3%, 3.0%) *t*=0.77	7.7% (0.5%, 14.9%) *t*=2.09
GLM *df* mtch	38	1.1% (−0.4%, 2.6%) *t*=1.39	7.3% (1.0%, 13.5%) *t*=2.27	0.8% (−1.5%, 3.1%) *t*=0.69	5.8% (−2.1%, 13.6%) *t*=1.44
GLM 4kn/yr	35	1.0% (−0.5%, 2.5%) *t*=1.35	7.6% (1.4%, 13.9%) *t*=2.40	0.7% (−1.6%, 3.0%) *t*=0.58	5.8% (−2.1%, 13.7%) *t*=1.45
GLM 12kn/yr	103	1.5% (−0.1%, 3.0%) *t*=1.78	7.9% (1.4%, 14.3%) *t*=2.37	0.4% (−2.0%, 2.9%) *t*=0.33	8.7% (0.3%, 17.0%) *t*=2.03
Portage					
SDN-GAM		Not included	Not included	Not included	Not included
KMHND-GAM		Not included	Not included	Not included	Not included
GAM 2002	36.9	1.1% (−0.8%, 2.9%) *t*=1.11	0.9% (−9.0%, 10.9%) *t*=0.19	1.9% (−1.0%, 4.9%) *t*=1.29	2.3% (−7.4%, 12.1%) *t*=0.47
GLM *df* mtch	38	−0.2% (−2.4%, 1.9%) *t*=−0.21	−1.2% (−12.4%, 10.1%) *t*=−0.20	0.9% (−2.4%, 4.3%) *t*=0.54	−1.5% (−12.5%, 9.5%) *t*=−0.27
GLM 4kn/yr	36	−0.0% (−2.2%, 2.1%) *t*=−0.05	−1.3% (−12.5%, 9.9%) *t*=−0.22	1.3% (−2.0%, 4.6%) *t*=0.76	−3.4% (−14.4%, 7.6%) *t*=−0.61
GLM 12kn/yr	106	−0.6% (−2.8%, 1.7%) *t*=−0.49	−2.3% (−14.4%, 9.8%) *t*=−0.38	0.8% (−2.7%, 4.3%) *t*=0.44	−1.1% (−12.9%, 10.7%) *t*=−0.18
Topeka					
SDN-GAM		Not included	Not included	Not included	Not included
KMHND-GAM		Not included	Not included	Not included	Not included
GAM 2002	37.3	0.8% (−2.9%, 4.5%) *t*=0.43	−3.1% (−19.0%, 12.8%) *t*=−0.38	−1.7% (−8.2%, 4.7%) *t*=−0.53	11.0% (−5.5%, 27.4%) *t*=1.31
GLM *df* mtch	38	0.3% (−4.0%, 4.6%) *t*=0.13	−4.4% (−22.5%, 13.7%) *t*=−0.47	−2.2% (−9.4%, 5.1%) *t*=−0.59	9.9% (−9.4%, 29.3%) *t*=1.00
GLM 4kn/yr	37	0.2% (−4.1%, 4.4%) *t*=0.08	−3.3% (−21.4%, 14.7%) *t*=−0.36	−2.0% (−9.2%, 5.2%) *t*=−0.55	10.7% (−8.5%, 30.0%) *t*=1.09
GLM 12kn/yr	108	−0.7% (−5.3%, 3.9%) *t*=−0.30	−6.6% (−26.0%, 12.8%) *t*=−0.67	−2.7% (−10.6%, 5.1%) *t*=−0.69	9.2% (−10.5%, 29.0%) *t*=0.91
Combined					
SDN-GAM		1.7% (1.2%, 2.2%) *t*=7.40	3.3% (1.0%, 5.7%) *t*=2.86	2.1% (1.4%, 2.8%) *t*=5.85	4.0% (1.8%, 6.2%) *t*=3.61
KMHND-GAM		1.6% (1.1%, 2.0%) *t*=6.75	2.0% (−0.3%, 4.2%) *t*=1.70	2.0% (1.3%, 2.7%) *t*=5.50	4.5% (2.2%, 6.7%) *t*=4.00
GAM 2002		1.4% (1.0%, 1.9%) *t*=6.37	2.3% (0.1%, 4.5%) *t*=2.04	1.8% (1.1%, 2.5%) *t*=5.18	4.1% (1.9%, 6.2%) *t*=3.74
GLM *df* mtch		1.0% (0.5%, 1.5%) *t*=3.81	1.6% (−0.9%, 4.0%) *t*=1.24	1.5% (0.7%, 2.3%) *t*=3.75	2.4% (−0.1%, 4.8%) *t*=1.92
GLM 4kn/yr		1.0% (0.5%, 1.5%) *t*=3.87	1.7% (−0.7%, 4.1%) *t*=1.36	1.5% (0.7%, 2.3%) *t*=3.70	2.7% (0.3%, 5.2%) *t*=2.21
GLM 12kn/yr		0.5% (−0.0%, 1.0%) *t*=1.79	1.6% (−1.0%, 4.2%) *t*=1.18	0.8% (0.0%, 1.6%) *t*=1.96	1.1% (−1.4%, 3.7%) *t*=0.86

[a] SDN = Schwartz et al 1996; KMHND = Klemm et al 2000; COPD = chronic obstructive pulmonary disease; IHD = ischemic heart disease.
SDN-GAM = GAM with LOESS for long-term trends and weather as published in SDN Table 5. SDN investigators provided previously unreported SO_4 results for the KMHND analysis.
KMHND-GAM = GAM with LOESS for long-term trends and weather as published in KMHND Table 7.
GAM 2002 = KMHDN-GAM with LOESS for long-term trends and weather but with the more stringent convergence criterion suggested by EPA (S-Plus string epsilon=10e-8,maxit=1000,bf.epsilon=10e-8,bf.maxit=1000,trace=T).
GLM *df* mtch = GLM with natural splines for weather and long-term trends using 38 *df* for time to match the approximate GAM 2002 51 *df* for each study area.
GLM 4kn/yr = GLM with natural splines for weather and long-term trends using a knot placed on the 21st of March, June, September and December (approximately every 90 days).
GLM 12kn/yr = GLM with natural splines for weather and long-term trends using a knot placed on the 21st of each month (approximately every 30 days).

[b] For $PM_{2.5}$; the degree of freedom for time varied slightly for other particulate mass concentrations.

Figure 1. Estimated increase in daily mortality associated with a 10 μg/m³ increase in PM$_{2.5}$ by degree of freedom for time: GAMs with LOESS.

Figure 3. Estimated increase in daily mortality associated with a 10 μg/m³ increase in PM$_{2.5}$ by degree of freedom for time: GLMs with natural spline smoothers.

Figure 2. Standard error of the estimated increase in daily mortality associated with a 10 μg/m³ increase in PM$_{2.5}$ by degree of freedom for time: GAMs with LOESS.
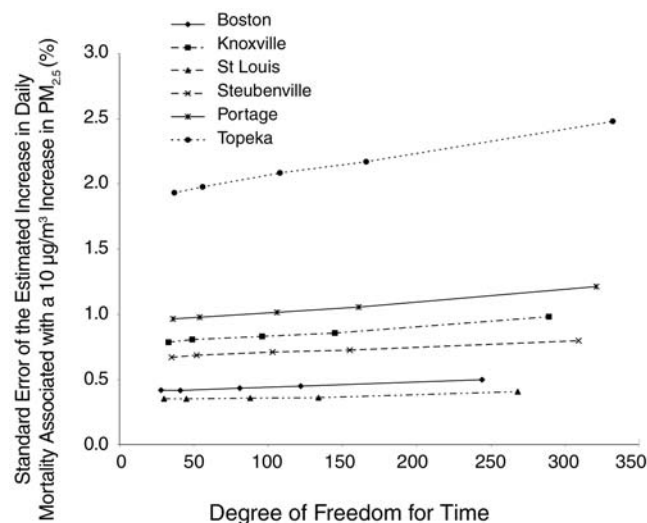
Figure 4. Standard error of the estimated increase in daily mortality associated with a 10 μg/m³ increase in PM$_{2.5}$ by degree of freedom for time: GLMs with natural spline smoothers.

Figure 3 shows that the estimated effects decrease monotonically only for Boston and Topeka. Figure 4 suggests that the degree of freedom should be limited.

DISCUSSION

The estimated effects calculated using the more stringent convergence criterion for GAMs with LOESS for long-term trends and weather were lower than those presented in KMHND and SDN. Estimated effects calculated with comparable GLMs using natural spline smoothers for long-term trends and weather were generally even lower. The relative magnitude of the estimated effect differed by study area and number of knots used to smooth time (Tables 1 and 2; Figure 3).

Figure 2 illustrates the variance (standard error) of the estimated effect underestimation problem discussed at the EPA Workshop on GAM-Related Statistical Issues in PM Epidemiology (November 4–6, 2002). The standard errors are flat and do not vary by the amount of long-term trend smoothing.

171

Comparison of the reduced estimated effects calculated using GLM with natural spline smoothers or GAM with LOESS suggests that further work and procedures are needed to evaluate the estimates of a functional relation estimated using different function methods. During the EPA Workshop, suggestions included the Akaike information criterion (AIC) and autocorrelation of the residuals. We did not compare AIC values because technical support staff of Insightful Corp (purveyors of S-Plus) verified a bug in both S-Plus 2000 and S-Plus 6.x regarding the AIC function (Insightful Corp, personal communication, 12/4/02). We did compare residual deviances, which are printed along with the S-Plus summary, and autocorrelations of the residuals, both of which were calculated without any apparent software bug.

The residual deviances from the GLM with 38 *df* for time were comparable, although slightly larger, than those obtained with the GAMs with the more stringent convergence criterion. The residual deviances were lower for GLMs using 12 knots per year compared to those using 4 knots per year to smooth time. For Boston, residual deviances in different models were as follows: KMHND-GAM, 2009.9; GAM 2002, 2003.8; GLM *df* mtch, 2027.9; GLM 4kn/yr, 2100.7; and GLM 12kn/yr, 1884.9. Autocorrelations of the residuals illustrated slight differences for each study area across the various methods of calculation. Negative autocorrelations predominated for lags up to 30 for all study areas using all methods of calculation. For Boston, 21 autocorrelations of lags 1 through 30 were negative for GAM 2002 and GLM 12kn/yr; the pattern was more uniform throughout the 30 GLM 12kn/yr lags than the GAM 2002 lags: 11 were negative among lags 1 through 15 and 10 were negative among lags 16 through 30 for GLM 12kn/yr compared with 8 among lags 1 through 15 and 13 among lags 15 through 30 for GAM 2002.

Combined estimates calculated using GAMs involved a greater weight for Boston than those calculated using GLM. This difference suggests that the GAM variance estimate may more grossly underestimate the variance of the estimated effects for Boston than for the other study areas. This aspect of GAM is consistent with the greater differential between the estimated effects calculated using GAM or GLM for Boston (Table 1) and suggests that more degrees of freedom may be required to smooth time in Boston (residual plots).

REFERENCES*

Schwartz J, Dockery DW, Neas LM. 1996. Is daily mortality associated specifically with fine particles? J Air Waste Manage Assoc 46:927–939.

Klemm RJ, Mason RM Jr, Heilig CM, Neas LM, Dockery DW. 2000. Is daily mortality associated specifically with fine particles? Data reconstruction and replication of analyses. J Air Waste Manage Assoc 50:1215–1222.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
CI	confidence interval
CM	coarse mass
COPD	chronic obstructive pulmonary disease
df	degree of freedom
EPA	Environmental Protection Agency (US)
GAM	generalized additive model
GLM	generalized linear model
IHD	ischemic heart disease
KMHND	Klemm et al 2000 article
LOESS	locally weighted smoothers
PM	particulate matter
PM_{10}	PM less than 10 µm in diameter
PM_{15}	PM less than 15 µm in diameter
$PM_{2.5}$	fine PM (PM less than 2.5 µm in diameter)
SDN	Schwartz et al 1996 article
SO_4	sulfate

* Bold type identifies publication containing the original analyses revised in this short communication report.

Short-Term Effects of Particulate Air Pollution on Cardiovascular Diseases in Eight European Cities

Le Tertre A, Medina S, Samoli E, Forsberg B, Michelozzi P, Boumghar A,
Vonk JM, Bellini A, Atkinson R, Ayres JG, Sunyer J, Schwartz J, and Katsouyanni K

ABSTRACT

The APHEA2 (Air Pollution and Health: a European Approach 2*) project investigated the short-term effects of air pollution on daily mortality and hospital admissions in 29 European cities. The original analyses, using LOESS (locally weighted smoothers) within a generalized additive model (GAM) framework, were conducted using S-Plus software with default convergence criteria settings. The sensitivity of the original results for cardiovascular admissions to more stringent convergence criteria were investigated, and a sensitivity analysis using alternative smoothing functions (natural splines and penalized splines with the same number of degrees of freedom) was carried out for cardiac admissions among persons more than 65 years of age. The results for all reinvestigated diseases and admission age group were insensitive to the specification of the convergence criteria. The results from the sensitivity analysis did not change significantly when penalized splines were applied but increased by approximately 50% when natural splines were applied.

The APHEA2 project has reported results on the short-term effects of ambient particle concentrations on both mortality and hospital admissions (Atkinson et al 2001; Katsouyanni et al 2001; Le Tertre et al 2002). Dominici et al (2002) identified a convergence problem in the standard S-Plus GAM procedure.

Responding to the request of the US Environmental Protection Agency, we present here the reanalysis of the data published in Le Tertre et al (2002) using more stringent convergence criteria in the GAM procedure. We also ran a sensitivity analysis on data for cardiac admissions among persons more than 65 years of age by using Poisson regression with natural splines or penalized splines to adjust for seasonal patterns and meteorologic variables.

DATA AND METHODS

We analyzed hospital admissions for cardiac diseases, ischemic heart diseases, and stroke by age group in eight European cities.

Particles were measured mostly by particulate matter less than 10 µm in diameter (PM_{10}), except in Paris (PM less than 13 µm in diameter [PM_{13}]) and in Milan and Rome (total suspended particulates, or TSP). To ensure comparability of the results, TSP estimates in Milan and Rome were scaled to equivalent PM_{10} measures using a locally derived conversion factor of 0.75. For Paris, the PM_{13} values were taken as being equivalent to PM_{10} in analyses.

Autoregressive Poisson models allowing for overdispersion were fitted. We controlled for potential confounders including long-term trend, season, day of the week, holidays, influenza epidemics, temperature and humidity. We used GAMs to adjust for nonlinear relations between confounders and morbidity using LOESS. Diagnostic tools for the models included the partial autocorrelation function of the residuals and the Akaïke information criterion (AIC).

We conducted a second-stage regression of local coefficients on effect modifiers that were not time dependent, following the method developed by Berkey et al (1998) More details on data and methods can be found in Le Tertre et al (2002).

For this reanalysis using more stringent criteria, we set the maximum number of iterations to 1000 and the difference of two successive coefficients to 10^{-14} instead of the default settings 10 and 10^{-3}, respectively.

We also applied, for cardiac admissions among persons more than 65 years of age, a sensitivity analysis using natural splines and penalized splines. These two approaches are fully parametric and were applied using the same number of degrees of freedom as those used in the original model.

RESULTS

Table 1 summarizes the effects of changing the convergence criteria used in the S-Plus software from the default values to the more stringent values. The PM_{10} effect estimates were largely unchanged when the more stringent criteria were used in the calculation of the regression estimates and their standard errors.

Figure 1 and Table 2 present the individual and combined estimates for cardiac admissions among persons more than 65 years of age for four scenarios: using the original criteria, using the more stringent criteria, using natural splines or using penalized splines instead of LOESS. The revised summary estimates increased by 6% using the stringent criteria, by 45% using natural splines, and by 8% using penalized splines over the estimates obtained in the original analysis.

Table 1. Results for PM_{10} Using Default and Revised Convergence Criteria. Pooled Percentage Increases for a 10 µg/m³ Increase in PM_{10}

		PM_{10}	
	Criteria	RR[a]	95% CI
Cardiac	Default	0.5	0.2, 0.8
	Revised	0.5	0.2, 0.8
Cardiac	Default	0.7	0.4, 1.0
over 65 yrs	Revised	0.7	0.4, 0.9
Ischemic heart	Default	0.3	−0.2, 0.6
disease below 65 yrs	Revised	0.3	−0.2, 0.6
Ischemic heart	Default	0.8	0.3, 1.2
disease over 65 yrs	Revised	0.7	0.3, 1.2
Stroke	Default	0.0	−0.3, 0.3
over 65 yrs	Revised	0.0	−0.3, 0.3

[a] RR = relative risk. Random- and fixed-effects models gave essentially the same results, as random-effects variance estimate was near zero.
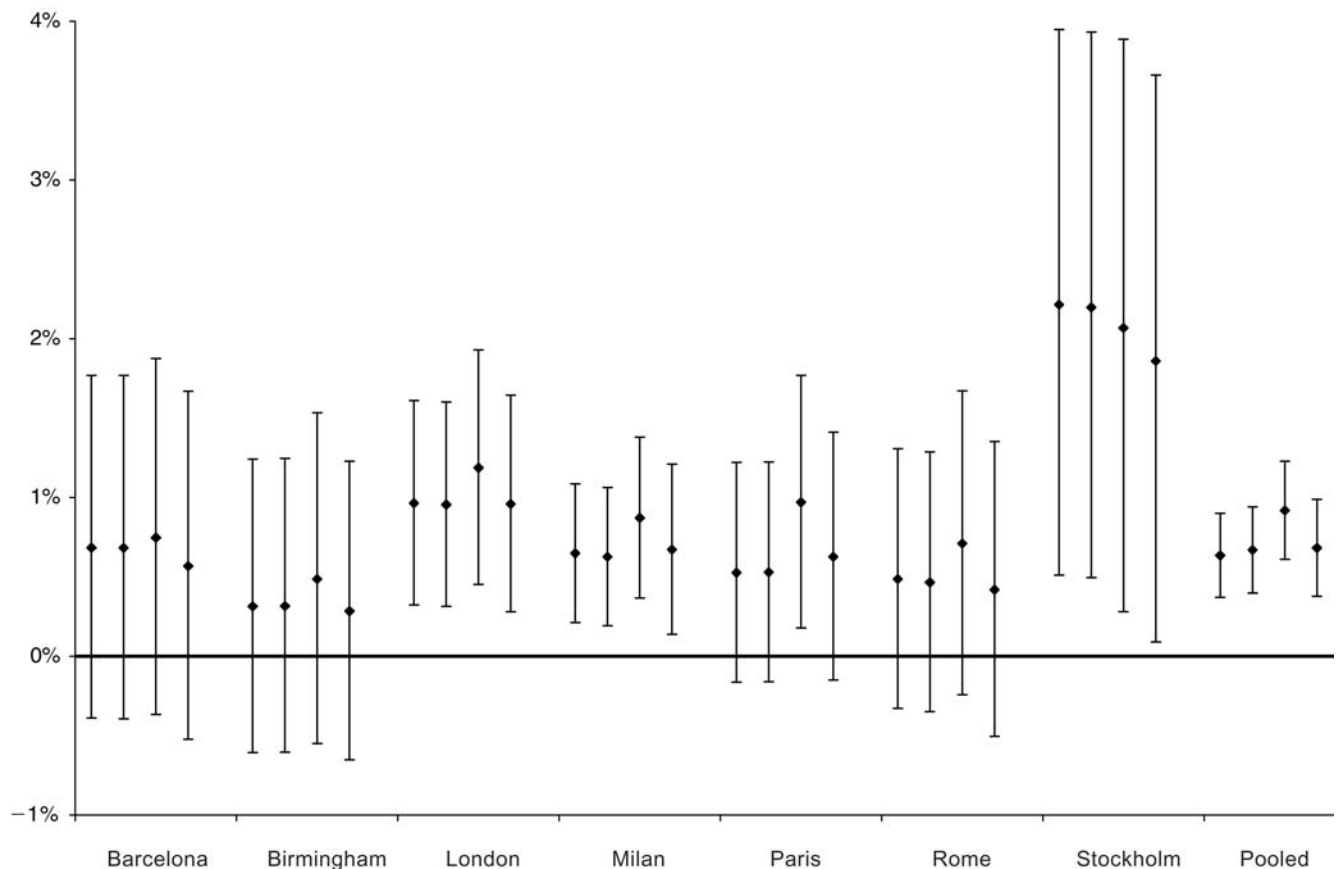
Figure 1. Sensitivity of results for PM_{10} and cardiac admissions among persons more than 65 years of age. Four results are shown for each city. Reading from left to right, the results from the original analysis are plotted followed by the results using the more stringent convergence criteria, using the natural spline smoothers and then the penalized spline smoothers, rather than LOESS functions. Each result represents the percentage change (and 95% CI) in the mean number of admissions associated with an increase in PM_{10} of 10 µg/m³.

Table 2. Sensitivity of City-Specific and Mean Effect Estimates for PM_{10} and Cardiac Admissions Among Persons More Than 65 Years of Age. Coefficients and Standard Errors (SEs) are Presented from the Original Analysis, Using More Stringent Convergence Criteria, Natural Spline Smoothers and Penalized Spline Smoothers

	LOESS Original Criteria		LOESS Stricter Criteria		GLM Using Natural Spline Smoothers		Penalized Splines	
	β	SE	β	SE	β	SE	β	SE
PM_{10}								
Barcelona	0.0006821	0.0005474	0.0006815	0.0005480	0.00074591	0.00056718	0.0005670	0.0005560
Birmingham	0.0003143	0.0004701	0.0003177	0.0004703	0.00048614	0.00052922	0.0002844	0.0004780
London	0.0009603	0.0003250	0.0009524	0.0003251	0.00118172	0.00037198	0.0009562	0.0003444
Milan	0.0006472	0.0002210	0.0006257	0.0002213	0.00086956	0.00025613	0.0006703	0.0002719
The Netherlands								
Paris	0.0005261	0.0003515	0.0005278	0.0003513	0.00096661	0.00040170	0.0006262	0.0003960
Rome	0.0004861	0.0004153	0.0004654	0.0004155	0.00070852	0.00048488	0.0004192	0.0004721
Stockholm	0.0021919	0.0008575	0.0021752	0.0008579	0.00204824	0.00090120	0.0018435	0.0008941
Pooled Fixed								
	0.0006333	0.0001342	0.0006685	0.0001374	0.00091537	0.0001565	0.00068102	0.00015489
Pooled Random								
	0.0006333	0.0001342	0.0006685	0.0001374	0.00091537	0.0001565	0.00068102	0.00015489
χ^2	0.56		0.56		0.80		0.76	

DISCUSSION

In this reanalysis, we showed that our results are robust to the convergence criteria used by the S-Plus software and, therefore, that the overall conclusions from the original study (Le Tertre et al 2002) remain unaltered. The results from the sensitivity analysis showed no significant change when penalized splines were used and an approximate 50% increase when natural splines were used. Natural splines are more sensitive to the number and location of the knots compared to penalized splines. In this reanalysis, we used the same number of degrees of freedom in all models. However, the smoothing parameter for a given number of degrees of freedom is not equivalent in each smoother. Hence the performance of each smoother for a given number of degrees of freedom will be different. For meaningful comparisons, the performance of each smoother under different scenarios needs to be evaluated further.

ACKNOWLEDGMENTS

REFRENCES*

Atkinson et al. 2001. Acute Effects of Particulate Air Pollution on Respiratory Admissions – Results from the APHEA2 Project. Am J Respir Crit Care Med 164:1860–1866.

Berkey CS, Hoaglin DC, Antczak-Bouckoms A, et al. Meta-analysis of multiple outcomes by regression with random effects. Stat Med 1998;17:2537–2550.

Dominici F, McDermott A, Zeger S, Samet J. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiol 2002; 156: 193–203.

Katsouyanni K et al. 2001. Confounding and Effect Modification in the Short-Term Effects of Ambient Particles on Total Mortality: Results from 29 European Cities within the APHEA2 Project. Epidemiology 12:521–531.

Le Tertre A, Medina S, Samoli E, Forsberg B, Michelozzi P, Boumghar A, Vonk JM, Bellini A, Atkinson R, Ayres JG, Sunyer J, Schwartz J, Katsouyanni K. 2002. Short-term effects of particulate air pollution on cardiovascular diseases in eight European cities. J Epidemiol Community Health 56:773–779.

* Bold type identifies publication containing the original analyses revised in this short communication report.

ABBREVIATIONS AND OTHER TERMS

APHEA2	Air Pollution and Health: A European Approach 2
CI	confidence interval
GAM	generalized additive model
IHD	ischemic heart disease

LOESS	locally weighted smoothers
PM_{10}	particulate matter less than 10 µm in diameter
PM_{13}	particulate matter less than 13 µm in diameter
RR	relative risk
TSP	total suspended particulates

Air Pollution and Cardiovascular Mortality in Phoenix, 1995–1997

Therese F Mar, Gary A Norris, Timothy V Larson, William E Wilson, and Jane Q Koenig

ABSTRACT

In this revised analysis of the association between air pollution and cardiovascular mortality in Phoenix, we compared the results for the generalized additive model (GAM*) with default convergence criteria (GAM-D), the GAM with strict convergence criteria (GAM-S), and the generalized linear model (GLM). In our original analysis using GAM-D, we found significant associations between cardiovascular mortality of elderly individuals (65–100 years of age) and various particulate matter (PM) metrics, source data from a factor analysis, and gaseous phase pollutants. With GAM-S and GLM we found similar associations. Our original findings remained unchanged. In Phoenix, elderly cardiovascular mortality was associated with various air pollutants: PM less than 10 μm in diameter (PM_{10}), PM less than 2.5 μm in diameter ($PM_{2.5}$), coarse fraction of PM, elemental carbon, motor vehicle exhaust, vegetative burning, regional sulfates, carbon monoxide (CO), nitrogen dioxide (NO_2), and sulfur dioxide (SO_2).

INTRODUCTION

This revised analysis of the association between air pollution and mortality in Phoenix, Arizona, from 1995 to 1997 is in response to the US Environmental Protection Agency (EPA)'s request that investigators who analyzed time series air pollution studies using GAMs reanalyze their data. This request is the result of recent concerns about the use of the default convergence criteria in S-Plus statistical software. In our original study (Mar et al 2000) we evaluated the association between mortality outcomes in elderly individuals and PM_{10}, $PM_{2.5}$, PMCF (PM_{10} minus $PM_{2.5}$), selected composition elements of $PM_{2.5}$, gaseous phase pollutants, and source data from a factor analysis. We found that cardiovascular mortality was

significantly associated with CO, NO_2, SO_2, $PM_{2.5}$, PM_{10}, PMCF, and elemental carbon. In this revised analysis we compared results from the original study that used GAM-D with GAM-S results and the GLM results.

METHODS

Mortality data for Maricopa County, Arizona, from 1995 to 1997 were obtained from the Arizona Center for Health Statistics in Phoenix. Only the data for residents 65 to 100 years of age in the zip code regions considered the most represented by the air pollution platform were included in this study. These zip code regions were recommended by the Arizona Department of Environmental Quality. PM_{10}, $PM_{2.5}$ and the chemical composition data for $PM_{2.5}$ were obtained from the EPA National Exposure Research Laboratory (NERL) platform in central Phoenix. Data for gaseous criteria pollutants (CO, NO_2 and SO_2) were obtained from EPA Aerometric Information Retrieval System (AIRS) database for residential sites in Phoenix.

STATISTICAL ANALYSIS

Poisson regression was used to evaluate the association between air pollution variables and cardiovascular mortality (*International Classification of Diseases* [ICD]-9 codes 390–448.9). All missing data were replaced with values averaged over the duration of the study. Base models were constructed by adjusting for time trends, temperature, and relative humidity with smoothing splines and for day of week with indicator variables. Degree of freedom (*df*) for time trend was determined by minimizing the autocorrelation in the residuals. For GAM-D and GAM-S, degree of freedom for the smoothing splines for temperature and relative humidity was determined by minimizing the Akaike information criterion (AIC). Lag days for temperature and relative humidity were similarly determined. With GLM we used natural spline smoothers for time trend, temperature, and relative humidity. As with GAM, degree of freedom and lags were chosen to minimize the AIC. For GAM-D, the convergence threshold was 0.001 and the maximum number of iterations was 10. With GAM-S the maximum iterations used was 1000, with the convergence criterion set to 10^{-8}. The trace

---

option was set to true to observe iteration details. With GLM, we used the default values for the number of iterations and convergence criterion. With GAM-D and GAM-S a smoothing spline with 10 df was used for time trends. Temperature was lagged one day with 2 df for the smoothing spline, and relative humidity was lagged 0 days with 2 df for the smoothing spline. With GLM, temperature was lagged 1 day with 2 df for the natural spline and relative humidity was lagged 0 days with 3 df for the natural spline. All relative risks (RRs) were calculated for an interquartile range increase in the air pollutant.

RESULTS

Table 1 compares the model estimates for GAM-D, GAM-S, and GLM. Factor 1 represents motor vehicle exhaust and resuspended road dust, factor 2 represents soil, factor 3 represents vegetative burning, factor 4 represents a local source of SO_2, and factor 5 represents regional sulfates.

In general, the β estimates using GAM-S were slightly smaller compared with those from GAM-D. The standard errors were consistent but negligibly larger. The β estimates from GLM were not consistently larger or smaller compared with GAM-D; however, the standard errors were consistently smaller. Table 2 compares the RRs for cardiovascular mortality associated with an interquartile range increase in each air pollutant and source factor. In the original GAM-D analysis, we had found associations between PM_{10}, $PM_{2.5}$, PMCF, elemental carbon, CO, NO_2 and SO_2 with cardiovascular mortality. The factors representing motor vehicles, vegetative burning and regional sulfates were also associated with cardiovascular mortality.

With GAM-S, we found similar associations. As with the GAM-D, PM_{10}, $PM_{2.5}$, and PMCF were associated with cardiovascular mortality; factors representing motor vehicles, vegetative burning, and regional sulfates were also significantly associated. In addition, elemental carbon, CO, NO_2 and SO_2 were associated with cardiovascular mortality. The strength of the association was nearly identical when the association obtained with GAM-D was significant ($P < 0.05$). With GLM compared with GAM-D, we also found very little or no difference in the RRs for cardiovascular mortality associated with each of the air pollution indicators. PM_{10} was marginally associated with cardiovascular mortality at 0 days lag (RR = 1.05 [95% confidence interval (CI): 1.00,1.09]) and 1 day lag (RR = 1.04 [1.00,1.08]). $PM_{2.5}$ was associated with cardiovascular mortality at 1 day lag (RR = 1.06 [1.01,1.11]) and PMCF was associated at 0 days lag (RR = 1.05 [1.01,1.09]). As with GAM-D and GAM-S, cardiovascular mortality was associated with elemental carbon and motor vehicle exhaust at 1 day lag, vegetative burning at 3 days lag, and regional sulfates at 0 days lag.

DISCUSSION

In this revised analysis of the Phoenix data using GAM-S and GLM compared with analysis using GAM-D, the RR for cardiovascular mortality associated with air pollution differed very little. When the association was significant, the choice of model did not affect the strength of the association. In the cases in which the association was only marginal using GAM-D, the use of GLM slightly reduced the strength of the association. The major conclusions concerning the association of air pollution and cardiovascular mortality in Phoenix, Arizona, from 1995 to 1997 remain unchanged. In Phoenix, the associations between various PM metrics, factor sources, and gaseous air pollutants with cardiovascular mortality in the elderly obtained using GAM-S and GLM are consistent with those associations found in the original analysis using GAM-D (Mar et al 2000).

REFERENCES*

Mar TF, Norris GA, Koenig JQ, Larson TV. 2000. Associations between air pollution and mortality in Phoenix, 1995–1997. Environ Health Perspect 108:347–353.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
CI	confidence interval
CO	carbon monoxide
df	degree of freedom
EPA	Environmental Protection Agency (US)
GAM	generalized additive model
GAM-D	GAM with default convergence criteria
GAM-S	GAM with strict convergence criteria
GLM	generalized linear model
NO_2	nitrogen dioxide
PM	particulate matter
$PM_{2.5}$	particulate matter less than 2.5 μm in diameter
PM_{10}	particulate matter less than 10 μm in diameter
PMCF	PM_{10} minus $PM_{2.5}$
RR	relative risk
SO_2	sulfur dioxide

* Bold text identifies publication containing the original analyses revised in this short communication report.

Table 1. Comparison of Model Estimates for GAM-D, GAM-S and GLM for Cardiovascular Mortality from Interquartile Range (IQR) Increase in Pollutants[a]

Pollutant	IQR	Lag	GAM-D β	GAM-D SE	GAM-S β	GAM-S SE	GLM β	GLM SE
PM$_{10}$	24.88	0	1.88E-03	7.66E-04	1.85E-03	7.73E-04	1.82E-03	8.70E-04
PM$_{10}$		1	1.47E-03	7.56E-04	1.43E-03	7.64E-04	1.49E-03	8.47E-04
PM$_{10}$		2	6.79E-04	7.60E-04	6.30E-04	7.68E-04	6.70E-04	8.53E-04
PM$_{10}$		3	1.08E-03	7.56E-04	1.04E-03	7.64E-04	1.11E-03	8.34E-04
PM$_{10}$		4	1.20E-03	7.57E-04	1.18E-03	7.64E-04	1.33E-03	8.26E-04
PM$_{2.5}$	8.52	0	3.91E-03	2.38E-03	3.71E-03	2.41E-03	3.97E-03	2.83E-03
PM$_{2.5}$		1	6.85E-03	2.36E-03	6.61E-03	2.39E-03	6.98E-03	2.78E-03
PM$_{2.5}$		2	1.83E-03	2.35E-03	1.47E-03	2.38E-03	1.06E-03	2.78E-03
PM$_{2.5}$		3	4.86E-03	2.35E-03	4.63E-03	2.37E-03	4.45E-03	2.75E-03
PM$_{2.5}$		4	5.43E-03	2.35E-03	5.23E-03	2.38E-03	5.20E-03	2.74E-03
PMCF	18.39	0	2.51E-03	9.88E-04	2.47E-03	9.97E-04	2.42E-03	1.08E-03
PMCF		1	1.62E-03	9.78E-04	1.59E-03	9.89E-04	1.66E-03	1.06E-03
PMCF		2	8.30E-04	9.87E-04	7.99E-04	9.97E-04	9.27E-04	1.07E-03
PMCF		3	9.34E-04	9.84E-04	9.11E-04	9.94E-04	1.07E-03	1.05E-03
PMCF		4	1.04E-03	9.85E-04	1.03E-03	9.94E-04	1.28E-03	1.04E-03
OC	2983.50	0	1.63E-06	6.98E-06	1.01E-06	7.05E-06	7.33E-07	8.39E-06
OC		1	1.46E-05	6.82E-06	1.38E-05	6.89E-06	1.34E-05	8.27E-06
OC		2	9.82E-06	6.84E-06	8.64E-06	6.92E-06	6.26E-06	8.33E-06
OC		3	1.39E-05	6.89E-06	1.32E-05	6.96E-06	1.13E-05	8.30E-06
OC		4	−1.88E-07	7.13E-06	−1.04E-06	7.20E-06	−3.32E-06	8.40E-06
EC	1168.25	0	−4.43E-06	1.88E-05	−6.43E-06	1.90E-05	−9.38E-06	2.22E-05
EC		1	4.40E-05	1.82E-05	4.19E-05	1.83E-05	3.94E-05	2.16E-05
EC		2	2.04E-05	1.82E-05	1.73E-05	1.84E-05	9.27E-06	2.19E-05
EC		3	2.65E-05	1.83E-05	2.43E-05	1.85E-05	1.75E-05	2.20E-05
EC		4	7.10E-06	1.87E-05	4.84E-06	1.89E-05	−2.41E-06	2.21E-05
TC	4187.75	0	5.33E-07	5.19E-06	3.17E-08	5.24E-06	−3.65E-07	6.24E-06
TC		1	1.15E-05	5.05E-06	1.09E-05	5.10E-06	1.05E-05	6.13E-06
TC		2	7.02E-06	5.07E-06	6.12E-06	5.12E-06	4.19E-06	6.18E-06
TC		3	9.71E-06	5.10E-06	9.12E-06	5.15E-06	7.58E-06	6.17E-06
TC		4	4.11E-07	5.26E-06	−2.39E-07	5.31E-06	−2.08E-06	6.23E-06
Factor 1[b]	1.11	0	9.27E-03	2.02E-02	8.38E-03	2.04E-02	6.20E-03	2.28E-02
Factor 1		1	5.06E-02	1.96E-02	4.97E-02	1.97E-02	4.66E-02	2.23E-02
Factor 1		2	2.28E-02	1.98E-02	2.11E-02	2.00E-02	1.41E-02	2.27E-02
Factor 1		3	1.15E-02	1.99E-02	9.97E-03	2.01E-02	1.98E-03	2.27E-02
Factor 1		4	1.92E-02	1.98E-02	1.79E-02	2.00E-02	9.63E-03	2.25E-02

Table continues next page

[a] OC = organic carbon; EC = elemental carbon; TC = total carbon.

[b] Factor 1 = motor vehicle exhaust and resuspended road dust.

[c] Factor 2 = soil.

[d] Factor 3 = vegetative burning.

[e] Factor 4 = local source of SO$_2$.

[f] Factor 5 = regional sulfates.

Table 1 (*Continued*). Comparison of Model Estimates for GAM-D, GAM-S and GLM for Cardiovascular Mortality from Interquartile Range (IQR) Increase in Pollutants[a]

Pollutant	IQR	Lag	GAM-D β	GAM-D SE	GAM-S β	GAM-S SE	GLM β	GLM SE
Factor 2[c]	1.26	0	−4.55E-03	2.09E-02	−5.03E-03	2.11E-02	−2.09E-03	2.40E-02
Factor 2		1	−3.48E-03	2.06E-02	−4.42E-03	2.08E-02	−1.47E-03	2.35E-02
Factor 2		2	−2.95E-02	2.07E-02	−3.02E-02	2.09E-02	−2.64E-02	2.34E-02
Factor 2		3	−2.44E-02	2.04E-02	−2.46E-02	2.06E-02	−1.88E-02	2.28E-02
Factor 2		4	−2.17E-02	2.02E-02	−2.14E-02	2.04E-02	−1.18E-02	2.26E-02
Factor 3[d]	1.02	0	−8.45E-03	1.89E-02	−8.92E-03	1.90E-02	−9.55E-04	2.02E-02
Factor 3		1	1.68E-02	1.83E-02	1.62E-02	1.85E-02	2.33E-02	1.97E-02
Factor 3		2	1.39E-02	1.83E-02	1.30E-02	1.85E-02	1.86E-02	1.96E-02
Factor 3		3	4.75E-02	1.78E-02	4.67E-02	1.80E-02	4.97E-02	1.89E-02
Factor 3		4	7.22E-03	1.86E-02	6.15E-03	1.88E-02	9.78E-03	1.95E-02
Factor 4[e]	1.09	0	−1.83E-02	2.14E-02	−1.95E-02	2.16E-02	−3.23E-02	2.41E-02
Factor 4		1	3.76E-03	2.08E-02	1.80E-03	2.10E-02	−1.21E-02	2.36E-02
Factor 4		2	−7.40E-03	2.08E-02	−9.24E-03	2.10E-02	−2.70E-02	2.39E-02
Factor 4		3	−4.63E-03	2.09E-02	−6.26E-03	2.11E-02	−2.21E-02	2.37E-02
Factor 4		4	−6.07E-04	2.10E-02	−1.76E-03	2.12E-02	−1.35E-02	2.32E-02
Factor 5[f]	1.38	0	4.06E-02	2.00E-02	4.12E-02	2.02E-02	4.34E-02	2.19E-02
Factor 5		1	1.13E-02	2.08E-02	1.16E-02	2.09E-02	1.10E-02	2.25E-02
Factor 5		2	1.68E-02	2.03E-02	1.73E-02	2.05E-02	2.03E-02	2.20E-02
Factor 5		3	8.07E-04	2.02E-02	1.62E-03	2.03E-02	5.44E-03	2.19E-02
Factor 5		4	−1.37E-03	1.99E-02	−6.36E-04	2.01E-02	5.31E-03	2.17E-02
CO	1.19	0	4.49E-02	2.14E-02	4.29E-02	2.16E-02	4.00E-02	2.75E-02
CO		1	7.66E-02	2.07E-02	7.31E-02	2.09E-02	7.00E-02	2.73E-02
CO		2	5.79E-02	2.00E-02	5.28E-02	2.02E-02	4.07E-02	2.73E-02
CO		3	5.32E-02	2.03E-02	5.07E-02	2.04E-02	3.82E-02	2.69E-02
CO		4	6.43E-02	2.06E-02	6.22E-02	2.08E-02	5.28E-02	2.64E-02
NO_2	0.02	0	2.58E+00	1.65E+00	2.42E+00	1.67E+00	1.86E+00	2.04E+00
NO_2		1	4.88E+00	1.59E+00	4.62E+00	1.60E+00	3.79E+00	2.03E+00
NO_2		2	2.53E+00	1.54E+00	2.16E+00	1.55E+00	5.24E-01	2.03E+00
NO_2		3	2.76E+00	1.55E+00	2.54E+00	1.57E+00	1.20E+00	1.98E+00
NO_2		4	5.74E+00	1.57E+00	5.64E+00	1.58E+00	4.86E+00	1.93E+00
SO_2	2.79	0	1.10E-02	9.13E-03	9.93E-03	9.25E-03	5.77E-03	1.10E-02
SO_2		1	1.28E-02	8.79E-03	1.10E-02	8.92E-03	5.51E-03	1.08E-02
SO_2		2	1.63E-02	8.64E-03	1.48E-02	8.74E-03	8.82E-03	1.07E-02
SO_2		3	1.85E-02	8.65E-03	1.75E-02	8.73E-03	1.25E-02	1.05E-02
SO_2		4	2.49E-02	8.58E-03	2.39E-02	8.67E-03	2.05E-02	1.02E-02

[a] OC = organic carbon; EC = elemental carbon; TC = total carbon.

[b] Factor 1 = motor vehicle exhaust and resuspended road dust.

[c] Factor 2 = soil.

[d] Factor 3 = vegetative burning.

[e] Factor 4 = local source of SO_2.

[f] Factor 5 = regional sulfates.

Table 2. Comparison of Relative Risks (RRs) for Cardiovascular Mortality from an Interquartile Range Increase in Pollutants Obtained by Applying GAM-D, GAM-S and GLM[a]

Pollutant	Lag	GAM-D			GAM-S			GLM		
		RR	LCL	UCL	RR	LCL	UCL	RR	LCL	UCL
PM_{10}	0	1.05	1.01	1.09	1.05	1.01	1.09	1.05	1.00	1.09
PM_{10}	1	1.04	1.00	1.08	1.04	1.00	1.08	1.04	1.00	1.08
PM_{10}	2	1.02	0.98	1.06	1.02	0.98	1.05	1.02	0.98	1.06
PM_{10}	3	1.03	0.99	1.07	1.03	0.99	1.07	1.03	0.99	1.07
PM_{10}	4	1.03	0.99	1.07	1.03	0.99	1.07	1.03	0.99	1.08
$PM_{2.5}$	0	1.03	0.99	1.08	1.03	0.99	1.07	1.03	0.99	1.08
$PM_{2.5}$	1	1.06	1.02	1.10	1.06	1.02	1.10	1.06	1.01	1.11
$PM_{2.5}$	2	1.02	0.98	1.06	1.01	0.97	1.05	1.01	0.96	1.06
$PM_{2.5}$	3	1.04	1.00	1.08	1.04	1.00	1.08	1.04	0.99	1.09
$PM_{2.5}$	4	1.05	1.01	1.09	1.05	1.00	1.09	1.05	1.00	1.09
PMCF	0	1.05	1.01	1.09	1.05	1.01	1.08	1.05	1.01	1.09
PMCF	1	1.03	0.99	1.07	1.03	0.99	1.07	1.03	0.99	1.07
PMCF	2	1.02	0.98	1.05	1.01	0.98	1.05	1.02	0.98	1.06
PMCF	3	1.02	0.98	1.05	1.02	0.98	1.05	1.02	0.98	1.06
PMCF	4	1.02	0.98	1.06	1.02	0.98	1.06	1.02	0.99	1.06
OC	0	1.00	0.96	1.05	1.00	0.96	1.05	1.00	0.95	1.05
OC	1	1.04	1.00	1.09	1.04	1.00	1.08	1.04	0.99	1.09
OC	2	1.03	0.99	1.07	1.03	0.99	1.07	1.02	0.97	1.07
OC	3	1.04	1.00	1.09	1.04	1.00	1.08	1.03	0.99	1.09
OC	4	1.00	0.96	1.04	1.00	0.96	1.04	0.99	0.94	1.04
EC	0	0.99	0.95	1.04	0.99	0.95	1.04	0.99	0.94	1.04
EC	1	1.05	1.01	1.10	1.05	1.01	1.10	1.05	1.00	1.10
EC	2	1.02	0.98	1.07	1.02	0.98	1.06	1.01	0.96	1.06
EC	3	1.03	0.99	1.08	1.03	0.99	1.07	1.02	0.97	1.07
EC	4	1.01	0.97	1.05	1.01	0.96	1.05	1.00	0.95	1.05
TC	0	1.00	0.96	1.05	1.00	0.96	1.04	1.00	0.95	1.05
TC	1	1.05	1.01	1.09	1.05	1.00	1.09	1.04	0.99	1.10
TC	2	1.03	0.99	1.07	1.03	0.98	1.07	1.02	0.97	1.07
TC	3	1.04	1.00	1.09	1.04	1.00	1.08	1.03	0.98	1.09
TC	4	1.00	0.96	1.05	1.00	0.96	1.04	0.99	0.94	1.04
Factor 1[b]	0	1.01	0.97	1.06	1.01	0.97	1.05	1.01	0.96	1.06
Factor 1	1	1.06	1.01	1.10	1.06	1.01	1.10	1.05	1.00	1.10
Factor 1	2	1.03	0.98	1.07	1.02	0.98	1.07	1.02	0.97	1.07
Factor 1	3	1.01	0.97	1.06	1.01	0.97	1.06	1.00	0.95	1.05
Factor 1	4	1.02	0.98	1.07	1.02	0.98	1.07	1.01	0.96	1.06

Table continues next page

[a] LCL = lower 95% confidence limit; UCL = upper 95% confidence limit; OC = organic carbon; EC = elemental carbon; TC = total carbon.

[b] Factor 1 = motor vehicle exhaust and resuspended road dust.

[c] Factor 2 = soil.

[d] Factor 3 = vegetative burning.

[e] Factor 4 = local source of SO_2.

[f] Factor 5 = regional sulfates.

Table 2 (*Continued*). Comparison of Relative Risks (RRs) for Cardiovascular Mortality from an Interquartile Range Increase in Pollutants Obtained by Applying GAM-D, GAM-S and GLM[a]

Pollutant	Lag	GAM-D			GAM-S			GLM		
		RR	LCL	UCL	RR	LCL	UCL	RR	LCL	UCL
Factor 2[c]	0	0.99	0.94	1.05	0.99	0.94	1.05	1.00	0.94	1.06
Factor 2	1	1.00	0.95	1.05	0.99	0.94	1.05	1.00	0.94	1.06
Factor 2	2	0.96	0.92	1.01	0.96	0.91	1.01	0.97	0.91	1.02
Factor 2	3	0.97	0.92	1.02	0.97	0.92	1.02	0.98	0.92	1.03
Factor 2	4	0.97	0.93	1.02	0.97	0.93	1.02	0.99	0.93	1.04
Factor 3[d]	0	0.99	0.95	1.03	0.99	0.95	1.03	1.00	0.96	1.04
Factor 3	1	1.02	0.98	1.06	1.02	0.98	1.06	1.02	0.98	1.07
Factor 3	2	1.01	0.98	1.05	1.01	0.98	1.05	1.02	0.98	1.06
Factor 3	3	1.05	1.01	1.09	1.05	1.01	1.09	1.05	1.01	1.09
Factor 3	4	1.01	0.97	1.05	1.01	0.97	1.04	1.01	0.97	1.05
Factor 4[e]	0	0.98	0.94	1.03	0.98	0.93	1.03	0.97	0.92	1.02
Factor 4	1	1.00	0.96	1.05	1.00	0.96	1.05	0.99	0.94	1.04
Factor 4	2	0.99	0.95	1.04	0.99	0.95	1.04	0.97	0.92	1.02
Factor 4	3	0.99	0.95	1.04	0.99	0.95	1.04	0.98	0.93	1.03
Factor 4	4	1.00	0.96	1.05	1.00	0.95	1.04	0.99	0.94	1.04
Factor 5[f]	0	1.06	1.00	1.12	1.06	1.00	1.12	1.06	1.00	1.13
Factor 5	1	1.02	0.96	1.07	1.02	0.96	1.08	1.02	0.96	1.08
Factor 5	2	1.02	0.97	1.08	1.02	0.97	1.08	1.03	0.97	1.09
Factor 5	3	1.00	0.95	1.06	1.00	0.95	1.06	1.01	0.95	1.07
Factor 5	4	1.00	0.95	1.05	1.00	0.95	1.05	1.01	0.95	1.07
CO	0	1.05	1.00	1.11	1.05	1.00	1.11	1.05	0.98	1.12
CO	1	1.10	1.04	1.15	1.09	1.04	1.14	1.09	1.02	1.16
CO	2	1.07	1.02	1.12	1.06	1.02	1.12	1.05	0.98	1.12
CO	3	1.07	1.02	1.12	1.06	1.01	1.11	1.05	0.98	1.11
CO	4	1.08	1.03	1.13	1.08	1.03	1.13	1.06	1.00	1.13
NO_2	0	1.04	0.99	1.09	1.04	0.99	1.09	1.03	0.97	1.09
NO_2	1	1.08	1.03	1.13	1.07	1.02	1.13	1.06	1.00	1.13
NO_2	2	1.04	0.99	1.09	1.03	0.99	1.08	1.01	0.95	1.07
NO_2	3	1.04	1.00	1.09	1.04	0.99	1.09	1.02	0.96	1.08
NO_2	4	1.09	1.04	1.14	1.09	1.04	1.14	1.08	1.02	1.14
SO_2	0	1.03	0.98	1.08	1.03	0.98	1.08	1.02	0.96	1.08
SO_2	1	1.04	0.99	1.09	1.03	0.98	1.08	1.02	0.96	1.08
SO_2	2	1.05	1.00	1.10	1.04	0.99	1.09	1.02	0.97	1.09
SO_2	3	1.05	1.00	1.10	1.05	1.00	1.10	1.04	0.98	1.10
SO_2	4	1.07	1.02	1.12	1.07	1.02	1.12	1.06	1.00	1.12

[a] LCL = lower 95% confidence limit; UCL = upper 95% confidence limit; OC = organic carbon; EC = elemental carbon; TC = total carbon.

[c] Factor 2 = soil.

[d] Factor 3 = vegetative burning.

[e] Factor 4 = local source of SO_2.

[f] Factor 5 = regional sulfates.

Air Pollution and Daily Deaths and Hospital Admissions in Los Angeles and Cook Counties

Suresh H Moolgavkar

ABSTRACT

I present reanalyses of associations between air pollution and daily deaths and hospital admissions in Los Angeles and Cook counties in the United States that were originally reported in a series of papers in 2000 (Moolgavkar 2000a,b,c). The principal reason for conducting these reanalyses was to assess the impact of using convergence criteria that are more stringent than the default criteria used in the S-Plus software package. I also report the results of generalized linear model (GLM*) analyses using natural splines with the same degree of freedom (*df*) as the smoothing splines I used in the generalized additive model (GAM) analyses. I found that changes in the convergence criteria and the use of GLM instead of GAM can, but does not always, substantially impact the results of the analyses and their interpretation. In particular, use of the more stringent convergence criteria led generally to a decrease in the estimate of the effect and an increase in its standard error. The GLMs estimated lower risks than the corresponding GAMs. As reported in the original paper, I found that in Los Angeles carbon monoxide (CO) is the best single index of air pollution associations with health endpoints, far better than the mass concentration of either PM_{10} or of $PM_{2.5}$ (particulate matter less than 10 or 2.5 µm in diameter). In Cook County the results were not so clear cut. However, any one of the gases was at least as good an index of air pollution effects on human health as was PM_{10}.

INTRODUCTION

In a series of papers (Moolgavkar 2000a,b,c) I presented the results of analyses of the association between air pollu-

tion and daily deaths and hospital admissions in three counties (Cook, Maricopa, and Los Angeles) in the United States using GAMs. In this short communication report I give the results of reanalyses of a subset of data in Cook and Los Angeles counties. Whereas in the original analyses I dealt with the association between air pollution and total nonaccidental and cause-specific deaths (due to cardiovascular disease [CVD], cerebrovascular disease, and chronic obstructive pulmonary disease [COPD]), these reanalyses did not consider cerebrovascular deaths. In addition, the original publications had considered cardiovascular, cerebrovascular and COPD daily hospital admissions. In these reanalyses I did not deal with cerebrovascular admissions; in the original analyses I had reported only weak and inconsistent associations between air pollution and cerebrovascular admissions and deaths.

The main goal of these reanalyses was to investigate the extent to which replacing default convergence criteria in the S-Plus software package with more stringent criteria altered the results of the original analyses. I compared these results with those obtained using GLMs. My original analyses of the data had used smoothing splines with 30 *df* to model temporal trends. I also repeated these analyses using 100 *df* splines to model temporal trends.

As described in more detail in the original papers (Moolgavkar 2000a,b,c), the original analyses (of both daily deaths and daily hospital admissions) covered the nine-year period from 1987 to 1995. The ICD-9 (*International Classification of Diseases, Ninth Edition*) codes used to define the various disease categories are also described in the original publications. The descriptive statistics for the daily deaths and hospital admissions along with those for air pollution and weather variables are presented in tabular form in the original publications. In Cook County I had daily monitoring information for PM_{10}, CO, sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and ozone (O_3). In Los Angeles County I had daily monitoring information for the gases but information only every sixth day for PM_{10} and $PM_{2.5}$. For this report, I analyzed the data for all available criteria pollutants with the exception of

---

ozone, because my original analyses had indicated that any effect of ozone was strongly seasonal. For all pollutants, I investigated lags of 0 to 5 days.

METHODS

As in the original analyses, I first analyzed the data using Poisson regression, allowing for overdispersion in a GAM. All models included an intercept term, indicator variables for day of week, and a spline smoother (30 and 100 *df*) for temporal trends. The strategy used for developing a basic model for weather-related covariates is described in the original publications (Moolgavkar 2000a,b,c). Once I had chosen this basic model, it was kept fixed for the analyses in which air pollution was considered. I describe in some detail the models used for analyses of total nonaccidental mortality in Cook and Los Angeles counties. Other mortality endpoints and the hospital admissions endpoints were analyzed similarly; contact the author for details. In my original publications I had run all models using the default convergence criteria for GAM in S-Plus. In this reanalysis I ran GAM with both convergence criteria for GAM set to 10^{-8}. In addition, I ran GLMs using natural splines with the same degree of freedom as in the GAM.

For analyses of total mortality in Los Angeles, I used models with an intercept term, smoothers for temporal trends (either 30 or 100 *df*), smoothers for temperature (lagged 1 day, 6 *df*), smoothers for relative humidity (same day, 6 *df*), indicators for day of week, and the pollutant entered linearly. I used the same model for total mortality in Cook County excepting temperature lagged 2 days and relative humidity lagged 3 days.

RESULTS

With two counties, 5 endpoints (3 mortality and 2 hospital admissions endpoints), 6 lag periods (0 to 5 days), 2 smoothing strategies for temporal trends (30 and 100 *df*), and 2 convergence criteria for GAM analyses, in addition to GLM analyses, I ran scores of regression models. I summarize the most important results in Tables 1 through 20 and discuss them briefly in the text. For the reader wishing to make comparisons with my earlier papers (Moolgavkar 2000a,b,c) I note that, in this short communication report, all results for PM_{10} are reported for increases in concentration of 10 µg/m^3, whereas in my earlier papers I had used 25 µg/m^3 as the unit of change. For the other pollutants I used the same units as I did in the original papers.

TOTAL NONACCIDENTAL MORTALITY

Table 1 shows the results for the association between PM and total mortality in Los Angeles for the distinct lags and analysis strategies. Table 2 gives the same information for CO and SO_2. Results for NO_2 were similar and are not shown here. Table 1 shows that the association between PM_{10} and mortality was strongest with a 2-day lag. This table also indicates that both the estimated size and significance of the effect decreased with increasing smoothing and stringency of the convergence criterion. The effect size and significance were smallest when natural splines with 100 *df* were used to smooth temporal trends. As judged by the Akaike information criterion (AIC), the GAM with 100 *df* fit best, followed by the natural splines model with 100 *df*. For $PM_{2.5}$, the strongest association was seen at a lag of 1 day. The pattern of effect sizes and significance was quite similar to that seen for PM_{10}. For fine particles ($PM_{2.5}$) as well, AIC indicated that the GAM with 100 *df* describes the data best followed by the GLM with 100 *df*. For $PM_{2.5}$ in particular, the results dramatically differed depending on the model used for analyses.

Table 2 shows the association between gases (CO and SO_2) and daily mortality in Los Angeles. These gases showed strong associations with mortality, particularly at lags of 0, 1, 2, and 3 days. The pattern with increasing smoothing and stringency of convergence criterion was similar to that seen with PM (Table 1), with the GLM with 100 *df* natural splines showing the smallest and least significant effects. The gases clearly showed substantially stronger associations with mortality than either PM_{10} or $PM_{2.5}$. This result is particularly interesting for SO_2 because concentrations of this gas are quite low in Los Angeles.

Table 3 shows results for two-pollutant analyses with PM_{10} or $PM_{2.5}$ and CO. Results are for a lag of 2 days for PM_{10} and 1 day for $PM_{2.5}$, because these are the lags that showed the strongest associations in single-pollutant analyses. The effect of CO clearly dominates that of PM, with the PM coefficients becoming small and insignificant in the two-pollutant analyses.

Table 4 shows some results of analyses in Cook County. The general pattern of results is similar to that in Los Angeles, although not as dramatic. The association of PM with mortality was strongest at lags 0 and 1. For a 2-day lag, the association seen with the GAM with 30 *df* smoothing for temporal trends was greatly attenuated with 100 *df* smoothing of temporal trends in both GAMs and GLMs. In 2-pollutant models (results not shown) with PM_{10} and one of the gases, the PM_{10} association remained robust and statistically significant for a 0-day lag, whereas the coefficient for each of the gases attenuated and became insignificant. For a 1-day lag, however, the coefficient for PM_{10} attenuated

Table 1. Total Nonaccidental Mortality in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily Nonaccidental Deaths Associated with Increases of 10 µg/m^3 in PM$_{10}$ and PM$_{2.5}$[a]

	Lag (days)					
	0	1	2	3	4	5
PM$_{10}$						
GAM-30 (10^{-3})	**0.05**	**0.11**	**0.48**	**−0.07**	**0.14**	**−0.41**
	0.26	0.56	2.55	−0.35	0.76	−2.18
GAM-30 (10^{-8})	**0.05**	**0.10**	**0.47**	**−0.07**	**0.14**	**−0.41**
	0.24	0.54	2.51	−0.37	0.75	−2.18
NS-30	**~0**	**0.07**	**0.45**	**−0.11**	**0.11**	**−0.40**
	−0.01	0.31	2.07	−0.50	0.53	−1.88
GAM-100 (10^{-3})	**0.10**	**−0.01**	**0.38**	**−0.26**	**0.14**	**−0.41**
	0.55	−0.07	2.27	−1.52	0.86	−2.47
GAM-100 (10^{-8})	**0.07**	**−0.03**	**0.36**	**−0.28**	**0.13**	**−0.41**
	0.37	−0.16	2.08	−1.67	0.80	−2.48
NS-100	**0.04**	**−0.03**	**0.34**	**−0.31**	**0.16**	**−0.37**
	0.18	−0.13	1.60	−1.44	0.79	−1.82
PM$_{2.5}$						
GAM-30 (10^{-3})	**0.55**	**0.60**	**0.39**	**−0.14**	**−0.08**	**−0.55**
	1.78	1.99	1.35	−0.46	−0.29	−1.82
GAM-30 (10^{-8})	**0.54**	**0.59**	**0.38**	**−0.15**	**−0.09**	**−0.56**
	1.74	1.96	1.31	−0.51	−0.32	−1.86
NS-30	**0.40**	**0.55**	**0.24**	**−0.28**	**−0.28**	**−0.72**
	1.06	1.50	0.70	−0.80	−0.85	−2.09
GAM-100 (10^{-3})	**0.42**	**0.13**	**−0.03**	**−0.42**	**−0.16**	**−0.73**
	1.47	0.48	−0.10	−1.62	−0.61	−2.69
GAM-100 (10^{-8})	**0.32**	**0.10**	**−0.09**	**−0.48**	**−0.21**	**−0.76**
	1.15	0.35	−0.33	−1.88	−0.81	−2.81
NS-100	**0.30**	**−0.01**	**−0.13**	**−0.56**	**−0.30**	**−0.72**
	0.81	−0.02	−0.41	−1.73	−0.96	−2.24

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-3}) = GAM with 30 *df* smoothing splines for temporal trends and default convergence criteria; GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-30 = GLM with 30 *df* natural splines for temporal trends. Other models are to be interpreted similarly. For example, GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria.

and became insignificant, whereas the coefficient for each the gases was robust and remained significant.

CARDIOVASCULAR MORTALITY

In my original analyses of CVD mortality in Los Angeles, which used GAMs with 30 *df* smoothers for temporal trends, I had reported the strongest and only statistically significant association with PM$_{10}$ at lag 2. Use of the more stringent convergence criteria altered the original estimates and *t*-statistics very little. Using 100 *df* smoothers for temporal trends did not substantially alter the results. The estimated coefficients for PM$_{10}$ with both GAMs and GLMs decreased somewhat and the associated *t*-statistics were somewhat lower. In the original analyses I found that PM$_{2.5}$ was significantly associated with CVD mortality at lags 0 and 1. Use of the more stringent convergence criteria

did not substantially alter either the estimated effects or the *t*-statistics. Use of 100 *df* smoothers in GAMs, however, resulted in a decrease in both the estimated effects and the associated *t*-statistics so that the results remained significant only at lag 0 and not at lag 1. With natural splines (100 *df*) in GLMs, on the other hand, although the estimated coefficients remained virtually unchanged, the standard errors increased about 15% to 20% so that the association with PM$_{2.5}$ was no longer significant at any lag. These results are summarized in Table 5.

Table 6 shows the association of CVD mortality with changes in concentrations of CO and SO$_2$. The results for NO$_2$ were similar but are not shown. Again, the decrease in the size and significance of the effects with increased stringency of convergence and with increased smoothing of temporal trends are readily apparent.

Table 2. Total Nonaccidental Mortality in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily Nonaccidental Deaths Associated with Increases of 1 ppm CO and 10 ppb SO_2[a]

	Lag (days)					
	0	1	2	3	4	5
CO						
GAM-30 (10^{-3})	**2.64**	**3.26**	**3.14**	**2.53**	**1.77**	**1.55**
	14.33	18.08	17.56	13.98	9.63	8.30
GAM-30 (10^{-8})	**2.60**	**3.19**	**3.06**	**2.48**	**1.73**	**1.51**
	13.98	17.64	16.94	13.52	9.27	8.03
NS-30	**1.94**	**2.55**	**2.30**	**1.65**	**0.90**	**0.72**
	7.78	10.25	9.01	6.52	3.62	2.90
GAM-100 (10^{-3})	**1.44**	**1.98**	**1.50**	**0.80**	**0.04**	**−0.06**
	8.63	12.07	9.22	4.85	0.26	−0.33
GAM-100 (10^{-8})	**1.17**	**1.64**	**1.07**	**0.39**	**−0.26**	**−0.35**
	7.02	9.98	6.55	2.39	−1.59	−2.08
NS-100	**1.11**	**1.56**	**0.97**	**0.31**	**−0.32**	**−0.39**
	4.59	6.38	3.81	1.22	−1.32	−1.66
SO_2						
GAM-30 (10^{-3})	**8.90**	**11.34**	**9.45**	**7.48**	**4.61**	**4.21**
	11.04	14.51	12.12	9.60	5.89	5.36
GAM-30 (10^{-8})	**8.75**	**11.58**	**9.32**	**7.37**	**4.53**	**4.14**
	10.87	14.28	11.95	9.47	5.79	5.28
NS-30	**6.96**	**9.05**	**6.93**	**5.23**	**2.70**	**2.51**
	7.10	9.33	7.11	5.49	2.88	2.71
GAM-100 (10^{-3})	**5.44**	**7.16**	**4.20**	**2.49**	**0.24**	**0.32**
	7.47	10.10	5.94	3.53	0.34	0.45
GAM-100 (10^{-8})	**4.72**	**6.08**	**3.11**	**1.68**	**−0.41**	**−0.17**
	6.48	8.57	4.40	2.38	−0.58	−0.24
NS-100	**4.53**	**5.83**	**2.83**	**1.47**	**−0.50**	**−0.22**
	4.82	6.16	2.94	1.59	−0.56	−0.25

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-3}) = GAM with 30 *df* smoothing splines for temporal trends and default convergence criteria; GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-30 = GLM with 30 *df* natural splines for temporal trends. Other models are to be interpreted similarly. For example, GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria.

In my original analyses (Moolgavkar 2000a,b,c) I had reported that in two-pollutant models with either PM_{10} or $PM_{2.5}$ and CO, the PM effect was attenuated, whereas the CO coefficient was robust and remained significant at all lags. This conclusion was not altered in the current analyses in models with 30 *df* smoothers for temporal trends, with CO clearly being the pollutant most strongly associated with CVD mortality. However, with 100 *df* for temporal trends, $PM_{2.5}$ appeared to show the strongest associations at lags 0 and 1 and CO at lag 2. These results are shown in Table 7.

In the original analyses of CVD mortality in Cook County, which used GAMs with 30 *df* smoothers for temporal trends, I had reported the strongest and the only statistically significant association with PM_{10} at lag 3 (estimated increase in daily deaths associated with 10 µg/m^3 increase in PM_{10} = 0.43%; *t*-statistic = 2.25). This estimate and the corresponding *t*-statistic were virtually unchanged when I used the more stringent convergence criteria. However, when I used 100 *df* smoothers for temporal trends, which by AIC yielded a better fit, the coefficient for a 3-day lag was greatly attenuated and became nonsignificant. The coefficient at lag 0 was the only statistically significant coefficient (estimated increase in daily deaths associated with 10 µg/m^3 increase in PM_{10}

Table 3. Two-Pollutant Analyses (CO and either PM_{10} or $PM_{2.5}$) of Total Nonaccidental Mortality in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily Nonaccidental Deaths Associated with Increases of 10 µg/m^3 PM_{10} and $PM_{2.5}$ and 1 ppm CO[a]

	PM_{10}-2	CO-2	$PM_{2.5}$-1	CO-1
GAM-30 (10^{-3})	**−0.33**	**3.88**	**−0.56**	**3.65**
	−1.50	8.48	−1.40	5.71
GAM-30 (10^{-8})	**−0.32**	**3.79**	**−0.53**	**3.50**
	−1.44	8.28	−1.33	5.48
GAM-100 (10^{-3})	**0.02**	**2.10**	**−0.47**	**2.16**
	0.10	5.00	−1.27	3.66
GAM-100 (10^{-8})	**0.08**	**1.50**	**−0.33**	**1.49**
	0.40	3.63	−0.90	2.54
NS-100	**0.09**	**1.40**	**−0.33**	**1.17**
	0.37	2.03	−0.77	1.28

[a] The integer following each pollutant indicates the lag used. For example, CO-2 indicates that CO was introduced into model with lag 2. For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-3}) = GAM with 30 *df* smoothing splines for temporal trends and default convergence criteria; GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends. Other models are to be interpreted similarly. For example, GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria.

= 0.39%; *t*-statistic = 2.32). When I used GLMs with 100 *df* natural splines for temporal trends, this result remained virtually unchanged. See Table 8 for these results.

The original results reported for CO remained virtually unchanged by use of the more stringent convergence criteria. In the original analyses, the strongest and only significant association with CO was seen at lag 3. With 100 *df* smoothing of temporal trends, whether using GAMs or GLMs, the coefficients for CO were attenuated and were not significant at any lag. In my original analyses I had reported statistically significant associations with NO_2 at lags 2, 3, and 4. Using the more stringent convergence criteria did not substantially change these results. However, with 100 *df* smoothers for temporal trends, the coefficients for NO_2 were generally attenuated and did not remain significant at any lag. In the original analyses I had reported significant associations of SO_2 with CVD mortality at lags 0, 1, 2, 3, and 4. With the more stringent convergence criteria these results remained unaltered. However, with 100 *df* smoothers for temporal trends with both GAMs and GLMs, the associations were attenuated and remained significant only at lags 0 and 1.

In joint pollutant models with PM_{10} and SO_2, with 100 *df* smoothers for temporal trends in both GAM and GLM, the coefficient for PM_{10} at lag 0 was substantially attenuated (estimated effect associated with a 10 µg/m^3 increase in concentration changed from 0.39% to 0.18%) and became insignificant (t = 0.94), whereas the coefficients for SO_2 remained virtually unchanged and statistically significant.

CHRONIC OBSTRUCTIVE PULMONARY DISEASE MORTALITY

In my original analyses I had reported that neither PM_{10} nor $PM_{2.5}$ appeared to be associated with COPD deaths in Los Angeles. The current analyses did not alter that conclusion. Although the estimated effects and their standard errors were affected by the stringency of the convergence criteria and the degree of smoothing of temporal trends (with the estimates of effect decreasing and the standard errors increasing with increasing smoothing and stringency of convergence criteria), the general pattern of results was similar: the estimated coefficients for PM_{10} were negative at lags 3 and 5, and the estimated coefficients for $PM_{2.5}$ were negative at lags 0, 2, 3, and 4. See Table 9 for these results.

In the original analyses, CO and SO_2 were reported to be strongly and significantly associated with COPD mortality at all lags from 0 to 5 days. This result was little altered by increasing the stringency of the convergence criteria. However, using 100 *df* smoothers for temporal trends led to a decrease in the coefficient sizes by about 50%. The effects of both CO and SO_2 continued to be statistically significant at lags 1, 2, and 3 but not at lags 0, 4, and 5. In the original analyses, NO_2 was reported to be strongly and significantly

Table 4. Total Nonaccidental Mortality in Cook County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily Nonaccidental Deaths Associated with Estimated Changes for Increases of 10 µg/m^3 PM$_{10}$, 1 ppm CO, and 10 ppb NO$_2$ and SO$_2$[a]

	Lag (days)					
	0	1	2	3	4	5
PM$_{10}$						
GAM-30 (10^{-8})	**0.48**	**0.43**	**0.26**	**0.17**	**0.15**	**−0.13**
	4.57	3.86	2.27	1.51	1.37	−1.23
GAM-100 (10^{-8})	**0.51**	**0.38**	**0.09**	**0.04**	**0.09**	**−0.17**
	5.11	3.65	0.73	0.33	0.90	−1.67
NS-100	**0.51**	**0.37**	**0.03**	**~0**	**0.07**	**−0.18**
	4.71	3.28	0.20	−0.04	0.67	−1.68
CO						
GAM-30 (10^{-8})	**0.06**	**0.25**	**0.11**	**0.12**	**0.10**	**0.06**
	1.28	5.08	2.29	2.53	2.07	1.26
GAM-100 (10^{-8})	**0.06**	**0.25**	**0.47**	**0.02**	**0.04**	**0.02**
	1.27	5.40	1.02	0.53	0.79	0.33
NS-100	**0.05**	**0.25**	**0.05**	**0.02**	**0.03**	**0.01**
	0.96	4.62	0.85	0.36	0.56	0.19
NO$_2$						
GAM-30 (10^{-8})	**0.38**	**1.13**	**0.66**	**0.75**	**0.56**	**0.17**
	1.55	4.58	2.66	3.04	2.29	0.69
GAM-100 (10^{-8})	**0.57**	**1.25**	**0.31**	**0.25**	**0.26**	**−0.05**
	2.45	5.31	1.28	1.05	1.13	−0.19
NS-100	**0.53**	**1.26**	**0.32**	**0.21**	**0.21**	**−0.12**
	2.01	4.80	1.13	0.74	0.77	−0.47
SO$_2$						
GAM-30 (10^{-8})	**1.73**	**2.41**	**1.78**	**1.72**	**1.46**	**0.34**
	3.06	4.27	3.13	3.03	2.63	0.61
GAM-100 (10^{-8})	**2.25**	**2.59**	**0.75**	**0.42**	**0.79**	**−0.18**
	4.21	4.83	1.39	0.78	1.51	−0.34
NS-100	**2.27**	**2.59**	**0.55**	**0.28**	**0.70**	**−0.35**
	3.81	4.38	0.85	0.43	1.18	−0.60

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-3}) = GAM with 30 *df* smoothing splines for temporal trends and default convergence criteria; GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-100 = GLM with 30 *df* natural splines for temporal trends. Other models are to be interpreted similarly. For example, GAM-100(10$^{-8)}$ = GAM with 100 *df* smoothing splines and the more stringent convergence criteria.

associated with COPD mortality at all lags from 0 to 3 days. This result was not altered with the more stringent convergence criteria (results not shown). With increased smoothing of temporal trends in GAMs, however, the coefficients were attenuated but remained significant at lags of 1 and 2 days. With GLM with 100 *df* natural splines for temporal trends the coefficients and their significance were dramatically attenuated. See Table 10 for these results.

In Cook County, my original analyses showed the strongest and only significant association between PM$_{10}$ and COPD mortality at lag 2. This result remained unchanged

with the more stringent convergence criteria with only modest attenuation of the PM$_{10}$ coefficient. With 100 *df* smoothing of temporal trends, the PM$_{10}$ coefficient was about 20% smaller and statistically insignificant (Table 11). However, as judged by AIC, the models with less smoothing (30 *df*) of temporal trends fit the data better.

In my original analyses I had reported little association between CO and COPD deaths in Cook County. All analyses for the present report confirmed that finding. For NO$_2$ I had reported the strongest and only significant association at lag 1. This finding was also confirmed in the current analyses

Table 5. Cardiovascular Disease (CVD) Mortality in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily CVD Deaths Associated with Increases of 10 µg/m^3 PM$_{10}$ and PM$_{2.5}$[a]

	Lag (days)					
	0	1	2	3	4	5
PM$_{10}$						
GAM-30 (10^{-3})	**0.33**	**0.28**	**0.89**	**−0.32**	**0.28**	**−0.49**
	1.17	0.98	3.10	−1.11	0.99	−1.75
GAM-30 (10^{-8})	**0.33**	**0.28**	**0.88**	**−0.32**	**0.28**	**−0.49**
	1.16	0.96	3.08	−1.13	0.99	−1.75
GAM-100 (10^{-3})	**0.41**	**0.20**	**0.80**	**−0.46**	**0.40**	**−0.54**
	1.55	0.72	2.97	−1.72	1.53	−2.10
GAM-100 (10^{-8})	**0.40**	**0.19**	**0.77**	**−0.47**	**0.40**	**−0.53**
	1.51	0.70	2.87	−1.77	1.53	−2.06
NS-100	**0.44**	**0.23**	**0.77**	**−0.47**	**0.47**	**−0.52**
	1.36	0.68	2.30	−1.40	1.44	−1.67
PM$_{2.5}$						
GAM-30 (10^{-3})	**0.99**	**1.03**	**0.78**	**−0.30**	**−0.09**	**−0.83**
	2.19	2.30	1.72	−0.64	−0.20	−1.83
GAM-30 (10^{-8})	**0.99**	**1.03**	**0.77**	**−0.31**	**−0.10**	**−0.84**
	2.17	2.30	1.69	−0.69	−0.21	−1.85
GAM-100 (10^{-3})	**0.99**	**0.79**	**0.44**	**−0.57**	**−0.07**	**−1.17**
	2.32	1.85	1.02	−1.33	−0.17	−2.78
GAM-100 (10^{-8})	**0.97**	**0.80**	**0.39**	**−0.63**	**−0.12**	**−1.21**
	2.29	1.88	0.91	−1.49	−0.28	−2.88
NS-100	**0.97**	**0.69**	**0.28**	**−0.64**	**−0.18**	**−1.27**
	1.92	1.34	0.53	−1.24	−0.35	−2.57

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-3}) = GAM with 30 *df* smoothing splines for temporal trends and default convergence criteria; GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-30 = GLM with 30 *df* natural splines for temporal trends. Other models are to be interpreted similarly. For example, GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria.

(results not shown). Similarly, the finding in the original analyses of a significant association with SO$_2$ only at lag 2 was also confirmed here. In joint pollutant analyses with ozone and PM$_{10}$ I had reported in the original paper that the PM$_{10}$ coefficient was attenuated and became insignificant, whereas the ozone coefficient remained significant. I did not consider ozone in the current analyses but in joint analyses with SO$_2$, both the PM$_{10}$ (estimate = 0.75, *t* = 1.30) and SO$_2$ (estimate = 4.42, *t* = 1.30) coefficients were attenuated and neither remained significant.

CARDIOVASCULAR ADMISSIONS

My original analyses in Los Angeles were conducted in 2 broad age groups, 20 to 64 years and 65+ years. The current analyses were restricted to the 65+ group. In the original analyses, I had reported that PM$_{10}$ was strongly and significantly associated with CVD admissions at lags 0, 1, and 2, and that PM$_{2.5}$ was significantly associated with CVD admissions at lags 0 and 1. These results were little altered by using the more stringent convergence criteria, except at lag 2 for PM$_{10}$ (Table 12). However, with increased smoothing of temporal trends, the coefficients for PM$_{10}$ were attenuated and remained significant only at lags 0 and 1 (Table 12). The coefficients for PM$_{2.5}$ were also considerably attenuated when 100 *df* smoothers were used for temporal trends. However, they continued to be significant, albeit with smaller *t*-statistics, at lags 0 and 1.

I had reported strong and significant associations with each of the gases except ozone in my original analyses. CO and SO$_2$ were significantly associated at all lags, whereas NO$_2$ was significantly associated at all lags except lag 4. Using the more stringent convergence criteria did not alter

Table 6. Cardiovascular Disease (CVD) Mortality in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily CVD Deaths Associated with Estimated Changes for Increases of 1 ppm CO and 10 ppb SO_2[a]

	Lag (days)					
	0	1	2	3	4	5
CO						
GAM-30 (10^{-3})	**3.47**	**3.93**	**4.08**	**3.76**	**2.91**	**2.63**
	12.87	14.66	15.39	14.13	10.70	9.52
GAM-30 (10^{-8})	**3.40**	**3.81**	**4.08**	**3.67**	**2.83**	**2.56**
	12.62	14.23	15.39	13.78	10.41	9.27
GAM-100 (10^{-3})	**1.78**	**1.97**	**1.88**	**1.62**	**0.76**	**0.57**
	6.99	7.83	7.54	6.48	2.98	2.21
GAM-100 (10^{-8})	**1.32**	**1.52**	**1.22**	**1.03**	**0.32**	**0.22**
	5.20	6.01	4.89	4.10	1.24	0.86
NS-100	**1.21**	**1.43**	**1.12**	**0.93**	**0.24**	**0.16**
	3.23	3.74	2.82	2.39	0.64	0.44
SO_2						
GAM-30 (10^{-3})	**12.40**	**14.06**	**13.02**	**11.21**	**7.33**	**7.36**
	10.84	12.23	11.34	9.78	6.35	6.32
GAM-30 (10^{-8})	**12.21**	**13.71**	**12.77**	**11.01**	**7.18**	**7.25**
	10.68	11.92	11.14	9.62	6.22	6.22
GAM-100 (10^{-3})	**6.93**	**7.84**	**6.42**	**5.16**	**1.78**	**2.54**
	6.46	7.24	5.95	4.79	1.64	2.32
GAM-100 (10^{-8})	**5.68**	**6.21**	**4.65**	**3.90**	**0.98**	**1.96**
	5.28	5.74	4.31	3.63	0.91	1.80
NS-100	**5.31**	**5.94**	**4.34**	**3.60**	**0.83**	**1.88**
	3.69	3.99	2.88	2.49	0.59	1.36

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-3}) = GAM with 30 *df* smoothing splines for temporal trends and default convergence criteria; GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-30 = GLM with 30 *df* natural splines for temporal trends. Other models are to be interpreted similarly. For example, GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria.

Table 7. Two-Pollutant Analyses (CO and either PM_{10} or $PM_{2.5}$) of Cardiovascular Disease (CVD) Mortality in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily CVD Deaths Associated with Increases of 10 µg/m³ PM_{10} and $PM_{2.5}$ and 1 ppm CO[a]

	PM_{10}-2	CO-2	$PM_{2.5}$	CO	$PM_{2.5}$-1	CO-1
GAM-100 (10^{-8})	**0.50**	**1.48**	**1.78**	**−3.01**	**0.91**	**−0.40**
	1.56	2.31	3.39	−3.60	1.72	−0.48
NS-100	**0.56**	**1.15**	**1.88**	**−3.44**	**0.91**	**−0.82**
	1.43	1.05	3.04	−2.51	1.42	−0.57

[a] The integer following each pollutant indicates the lag used. For example, CO-2 indicates that CO was introduced into model with lag 2. For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

Table 8. Cardiovascular Disease (CVD) Mortality in Cook County: Estimated Percentage Change (log relative risk × 100) in Daily CVD Deaths Associated with Estimated Changes for Increases of 10 μg/m³ PM_{10} and 10 ppb SO_2[a]

	Lag (days)					
	0	1	2	3	4	5
PM_{10}						
GAM-30 (10^{-8})	**0.32**	**0.24**	**0.29**	**0.43**	**0.26**	**−0.07**
	1.75	1.28	1.49	2.25	1.41	−0.38
GAM-100 (10^{-8})	**0.39**	**0.21**	**0.07**	**0.30**	**0.22**	**−0.08**
	2.32	1.22	0.41	1.67	1.26	−0.49
NS-100	**0.39**	**0.21**	**−0.01**	**0.23**	**0.19**	**−0.10**
	2.22	1.14	−0.06	1.18	1.07	−0.57
SO_2						
GAM-30 (10^{-8})	**2.14**	**3.10**	**1.86**	**1.47**	**2.10**	**0.45**
	2.42	3.51	2.08	1.64	2.42	0.51
GAM-100 (10^{-8})	**2.52**	**3.03**	**0.46**	**−0.06**	**1.39**	**0.06**
	2.85	3.43	0.51	−0.07	1.60	0.07
NS-100	**2.44**	**2.90**	**~0**	**−0.32**	**1.23**	**−0.27**
	2.52	3.01	−0.003	−0.30	1.27	−0.29

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends; GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria.

Table 9. Chronic Obstructive Pulmonary Disease (COPD) Mortality in Los Angeles County: Estimated Percentage Change (log relative risk × 100) in Daily COPD Deaths Associated With Increases of 10 μg/m³ PM_{10} and $PM_{2.5}$[a]

	Lag (days)					
	0	1	2	3	4	5
PM_{10}						
GAM-30 (10^{-8})	**0.33**	**0.86**	**0.60**	**−1.88**	**0.27**	**−0.34**
	0.42	1.11	0.75	−2.37	0.34	−0.42
GAM-100 (10^{-8})	**0.26**	**1.21**	**0.30**	**−2.40**	**0.06**	**−0.20**
	0.34	1.57	0.38	−3.18	0.07	−0.26
NS-100	**0.37**	**1.21**	**0.24**	**−2.45**	**0.21**	**−0.14**
	0.39	1.25	0.24	−2.70	0.22	−0.15
$PM_{2.5}$						
GAM-30 (10^{-8})	**−0.56**	**0.38**	**−0.27**	**−3.47**	**−1.91**	**0.15**
	−0.45	0.30	−0.22	−2.84	−1.50	0.12
GAM-100 (10^{-8})	**−1.42**	**0.86**	**−0.49**	**−4.45**	**−3.21**	**0.55**
	−1.17	0.69	−0.39	−3.83	−2.58	0.44
NS-100	**−1.21**	**0.20**	**−0.62**	**−4.40**	**−3.15**	**0.83**
	−0.83	0.13	−0.41	−3.22	−2.13	0.56

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; GAM-100($^{-8}$) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria NS-100 = GLM with 100 *df* natural splines for temporal trends.

Table 10. Chronic Obstructive Pulmonary Disease (COPD) Mortality in Los Angeles County: Estimated Percentage Change (log relative risk × 100) in Daily COPD Deaths Associated with Estimated Changes for Increases of 1 ppm CO, and 10 ppb SO_2[a]

	Lag (days)					
	0	1	2	3	4	5
CO						
GAM-30 (10^{-8})	**3.32**	**4.60**	**5.63**	**5.48**	**3.90**	**3.72**
	4.38	6.13	7.62	7.39	5.17	4.85
GAM-100 (10^{-8})	**0.79**	**2.25**	**2.92**	**3.05**	**1.36**	**1.25**
	1.07	3.09	3.99	4.13	1.82	1.66
NS-100	**−0.72**	**0.24**	**1.17**	**1.43**	**−0.04**	**−0.04**
	−0.65	0.21	1.04	1.32	−0.04	−0.04
SO_2						
GAM-30 (10^{-8})	**13.35**	**18.99**	**18.37**	**17.66**	**13.48**	**11.88**
	4.11	5.85	5.68	5.48	4.16	3.63
GAM-100 (10^{-8})	**5.03**	**10.41**	**8.35**	**8.74**	**5.48**	**4.36**
	1.57	3.27	2.62	2.74	1.71	1.36
NS-100	**−0.68**	**4.58**	**3.98**	**6.05**	**3.51**	**2.56**
	−0.16	1.05	0.94	1.48	0.87	0.64

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

Table 11. Chronic Obstructive Pulmonary Disease (COPD) Mortality in Cook County: Estimated Percentage Change (log relative risk × 100) in Daily COPD Deaths Associated with Estimated Changes for Increases of 10 μg/m^3 PM_{10}, and 10 ppb SO_2[a]

	Lag (days)					
	0	1	2	3	4	5
PM_{10}						
GAM-30 (10^{-8})	**0.48**	**0.95**	**1.07**	**−0.03**	**0.59**	**−0.13**
	0.96	1.89	2.04	−0.06	1.89	−0.26
GAM-100 (10^{-8})	**0.48**	**0.88**	**0.88**	**−0.30**	**0.51**	**−0.24**
	1.00	1.76	1.66	−0.55	1.02	−0.47
NS-100	**0.55**	**0.94**	**0.89**	**−0.31**	**0.52**	**−0.20**
	1.05	1.75	1.45	−0.54	0.99	−0.38
SO_2						
GAM-30 (10^{-8})	**−2.08**	**2.06**	**5.49**	**4.74**	**1.32**	**−3.20**
	−0.80	0.79	2.11	1.82	0.51	−1.23
GAM-100 (10^{-8})	**−0.45**	**3.23**	**5.33**	**3.94**	**0.91**	**−3.64**
	−0.17	1.26	2.08	1.53	0.36	−1.42
NS-100	**−0.52**	**3.20**	**5.10**	**3.71**	**0.58**	**−4.17**
	−0.18	1.12	1.68	1.19	0.20	−1.46

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

Table 12. Hospital Admissions for Cardiovascular Disease (CVD) in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily CVD Admissions Associated with Increases of 10 µg/m^3 PM$_{10}$ and PM$_{2.5}$[a]

	Lag (days)					
	0	1	2	3	4	5
PM$_{10}$						
GAM-30 (10^{-8})	**0.67**	**0.51**	**0.33**	**0.02**	**−0.05**	**−0.09**
	3.06	2.30	1.37	0.07	−0.21	−0.39
GAM-100 (10^{-8})	**0.54**	**0.48**	**~0**	**−0.16**	**−0.26**	**−0.35**
	2.53	2.23	−0.06	−0.74	−1.20	−1.57
NS-100	**0.55**	**0.49**	**−0.07**	**−0.15**	**−0.28**	**−0.41**
	2.06	1.86	−0.25	−0.57	−1.07	−1.54
PM$_{2.5}$						
GAM-30 (10^{-8})	**1.58**	**1.39**	**~0**	**−0.27**	**0.29**	**0.36**
	4.59	3.88	0.12	−0.77	0.83	1.02
GAM-100 (10^{-8})	**1.16**	**1.13**	**−0.53**	**−0.59**	**0.09**	**0.02**
	3.46	3.32	−1.44	−1.75	0.26	0.07
NS-100	**1.26**	**1.20**	**−0.64**	**−0.56**	**−0.03**	**−0.03**
	3.04	2.88	−1.42	−1.36	−0.08	−0.07

[a] Admissions of persons $\geq$ 65 years of age. For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

these results. However, increased smoothing of temporal trends did. The coefficients for CO were considerably attenuated and remained significant at only lags 0, 1, and 2. The coefficients for SO$_2$ were also considerably attenuated and remained significant only at lags 0 and 1. See Table 13 for these results. Similar results were noted for NO$_2$ (results not shown).

As in the original analyses, joint analyses of either PM$_{10}$ or PM$_{2.5}$ with CO resulted in considerable attenuation of the PM coefficients, which were rendered insignificant, whereas the CO coefficients remained robust and significant (Table 14).

In my original analyses I had reported significant associations in Cook County between PM$_{10}$ and CVD admissions at lags 0, 1 and 2. In the current analyses, both the estimated effects and the *t*-statistics were somewhat smaller with the more stringent convergence criteria, particularly at lag 2. The same phenomenon was seen with 100 *df* smoothers for temporal trends. According to AIC, the latter models describe the data better than models with 30 *df* for temporal trends. The effect of using the more stringent convergence criteria was more striking: the PM$_{10}$ coefficient at lag 2 was reduced by almost 50% and lost statistical significance (Table 15).

In my original analyses I had reported strong associations between each of the gases and CVD admissions (CO at lags 0, 1, 2, 3, and 4; NO$_2$ at lags 0, 1, 2, and 3; SO$_2$ at lags

0, 1, 2, and 3). Using the more stringent convergence criteria did not greatly affect either the size or the significance of these associations. Using 100 *df* for temporal trends had more of an effect: the coefficients for each of the gases were attenuated and lost significance for some lags. In particular, the CO coefficients were attenuated and lost significance at lag 3 (Table 15). Similar results were seen for NO$_2$ and SO$_2$ (results not shown).

In joint analyses with PM$_{10}$ and one of the gases, both PM$_{10}$ and the gas remained significant at lag 0. At lag 1 the coefficient for the gas remained significant, whereas that for PM$_{10}$ did not (only CO data shown; see Table 16).

CHRONIC OBSTRUCTIVE PULMONARY DISEASE ADMISSIONS

In my original paper on COPD admissions in Los Angeles county (Moolgavkar 2000b) I had reported results of analyses conducted in three broad age groups, 0 to 19 years, 20 to 64 years, and 65+ years, as well as the results for total admissions in all age groups combined. For the present report I reanalyzed only total admissions over all age groups. In the original analyses I had reported statistically significant associations with PM$_{10}$ at lags of 0, 1, 2, and 3 and with PM$_{2.5}$ at lags of 0, 1, and 2. Using more stringent convergence criteria decreased the coefficients and the *t*-statistics somewhat but did not alter these conclusions. Going to 100 *df* to smooth temporal trends likewise altered

Table 13. Hospital Admissions for Cardiovascular Disease (CVD) in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily CVD Admissions Associated with Increases of 1 ppm CO and 10 ppb SO_2[a]

	Lag (days)					
	0	1	2	3	4	5
CO						
GAM-30 (10^{-8})	**4.88**	**3.38**	**2.18**	**1.21**	**1.04**	**1.11**
	23.76	15.81	10.02	5.53	4.71	4.99
GAM-100 (10^{-8})	**4.56**	**2.47**	**0.61**	**−0.58**	**−0.50**	**−0.23**
	22.80	11.86	2.87	−2.73	−2.34	−1.07
NS-100	**4.46**	**2.31**	**0.40**	**−0.77**	**−0.67**	**−0.38**
	15.00	7.30	1.19	−2.37	−2.14	−1.22
SO_2						
GAM-30 (10^{-8})	**15.58**	**10.25**	**5.44**	**1.57**	**1.19**	**1.21**
	17.44	11.05	5.76	1.66	1.25	1.27
GAM-100 (10^{-8})	**13.94**	**6.85**	**0.71**	**−2.94**	**−2.27**	**−1.60**
	16.06	7.60	0.78	−3.22	−2.48	−1.75
NS-100	**13.67**	**6.44**	**0.23**	**−3.37**	**−2.66**	**−1.91**
	11.82	5.23	0.18	−2.79	−2.26	−1.66

[a] Admissions of persons $\geq$ 65 years of age. For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

Table 14. Two-Pollutant Analyses (CO and either PM_{10} or $PM_{2.5}$) of Hospital Admissions for Cardiovascular Disease (CVD) in Los Angeles County: For Each Model, Estimated Percentage Change (log relative risk $\times$ 100) in Daily CVD Admissions Associated with Increases of 10 µg/m³ PM_{10} and $PM_{2.5}$ and 1 ppm CO[a]

	PM_{10}-0	CO-0	$PM_{2.5}$-0	CO-0	$PM_{2.5}$-1	CO-1
GAM-100 (10^{-8})	**−0.26**	**3.89**	**0.39**	**4.25**	**0.24**	**4.16**
	−1.02	7.58	0.92	6.24	0.53	5.86
NS-100	**−0.22**	**3.79**	**0.58**	**2.66**	**0.27**	**3.67**
	−0.70	4.66	1.15	2.37	0.52	3.16

[a] Admissions of persons $\geq$ 65 years of age. The integer following each pollutant indicates the lag used. For example, CO-0 indicates that CO was introduced into the model with lag 0. For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

the PM_{10} results very little but altered the results for $PM_{2.5}$ more, decreasing the estimates of effect and the *t*-statistics but leaving the basic qualitative conclusions unchanged (Table 17).

In my original analyses I had reported that each of the gases (CO, NO_2, and SO_2) was strongly and significantly associated with total COPD admissions for all lags between 0 and 5. These results continued to hold for all the current analyses, with the single exception that the coefficient for SO_2 became insignificant at lag 5 when 100 *df* smoother was used for temporal trends (Table 18; SO_2 results not shown).

In two-pollutant models with either PM_{10} or $PM_{2.5}$ and one of the gases I had reported in my original analyses that

the effects of the gases was robust, whereas that of PM was attenuated and became insignificant at all lags. In the reanalyses the results were not so clear cut. With the more stringent convergence criteria, the coefficients and *t*-statistics decreased somewhat, but the major difference was seen with models that employed more smoothing of temporal trends. With PM_{10} and either CO or SO_2 (results not shown), the coefficient for PM_{10} was significant at lags 1 and 2 and the coefficient for the gases was not. However, exactly the reverse was true for lags 0 and 3, with the coefficients for CO and SO_2 remaining significant, whereas the coefficients for PM_{10} were attenuated and became insignificant. With NO_2, however, the coefficient for PM_{10} was not

Table 15. Cardiovascular Disease (CVD) Admissions in Cook County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily CVD Admissions Associated with Estimated Changes for Increases of 10 $\mu g/m^3$ PM_{10}, and 1 ppm CO [a]

	Lag (days)					
	0	1	2	3	4	5
PM_{10}						
GAM-100 (10^{-3})	**0.76**	**0.51**	**0.26**	**−0.05**	**−0.10**	**−0.31**
	6.49	4.07	2.00	−0.41	−0.81	−2.57
GAM-100 (10^{-8})	**0.81**	**0.49**	**0.14**	**−0.11**	**−0.09**	**−0.29**
	6.89	3.95	1.06	−0.91	−0.78	−2.35
NS-100	**0.85**	**0.51**	**0.13**	**−0.12**	**−0.08**	**−0.26**
	6.76	3.77	0.83	−0.84	−0.57	−2.00
CO						
GAM-100 (10^{-3})	**4.74**	**3.47**	**1.61**	**0.84**	**0.07**	**−0.90**
	8.68	6.30	2.88	1.51	1.27	−1.57
GAM-100 (10^{-8})	**4.70**	**3.40**	**1.22**	**0.10**	**0.19**	**−1.23**
	8.68	6.17	2.18	0.18	0.34	−2.24
NS-100	**4.90**	**3.59**	**1.27**	**0.08**	**0.21**	**−1.14**
	7.70	5.51	1.90	0.11	0.32	−1.74

[a] Admissions of persons $\geq$ 65 years of age. For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-100(10^{-3}) = GAM with 100 *df* smoothing splines for temporal trends and the default convergence criteria; GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

Table 16. Two-Pollutant Analyses (CO and PM_{10}) of Hospital Admissions for Cardiovascular Disease (CVD) in Cook County: For Each Model, Estimated Percentage Change (log relative risk $\times$ 100) in Daily CVD Admissions Associated with Increases of 10 $\mu g/m^3$ PM_{10} and 1 ppm CO[a]

	PM_{10}-0	CO-0	PM_{10}-1	CO-1	PM_{10}-2	CO-2
GAM-100 (10^{-8})	**0.59**	**3.12**	**0.25**	**3.38**	**0.04**	**1.30**
	4.69	5.24	1.88	5.47	0.29	2.08
NS-100	**0.62**	**3.13**	**0.26**	**3.44**	**0.03**	**1.32**
	4.61	4.41	1.78	4.65	0.16	1.77

[a] Admissions of persons $\geq$ 65 years of age. The integer following each pollutant indicates the lag used. For example, CO-0 indicates that CO was introduced into the model with lag 0. For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines for temporal trends and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

significant at any lag, whereas the coefficient for NO_2 remained significant at lags 0, 1, and 3 (Table 19). With $PM_{2.5}$ and either CO or SO_2, the coefficient for $PM_{2.5}$ remained significant at lag 0. However, in joint analyses of $PM_{2.5}$ and NO_2, the coefficient for NO_2 was significant at lag 0, whereas that for $PM_{2.5}$ was attenuated and became insignificant (Table 19).

In Cook County I had information on COPD admissions only among the elderly (aged 65 years or more) and my analyses were, therefore, restricted to this age group. In my original analyses I had reported small and insignificant associations between COPD admissions and PM_{10} at lags 0 and 5. These results were somewhat attenuated with the more stringent convergence criteria (Table 20). With 100 *df*

smoothers for temporal trends, PM_{10} was significantly associated with admissions at lag 0, and there was some indication of an association at lag 1 (which was not significant). The coefficient at lag 5 became negative, however.

In the original publication I had reported a statistically significant association with CO at lag 1 and a near-significant association at lag 0. These conclusions continued to hold in the reanalyses with the more stringent convergence criteria and also with increased smoothing of temporal trends, although the coefficient and *t*-statistics were somewhat smaller. In the original analyses, NO_2 was significantly associated with COPD admissions at lags 0 and 5. In the current analyses, the results remained unchanged with the more stringent convergence criteria (Table 20). With

195

Table 17. Hospital Admissions for Chronic Obstructive Pulmonary Disease (COPD) in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily COPD Admissions Associated with Increases of 10 µg/m^3 PM$_{10}$ and PM$_{2.5}$[a]

	Lag (days)					
	0	1	2	3	4	5
PM$_{10}$						
GAM-30 (10^{-8})	**1.09**	**1.11**	**1.51**	**0.81**	**−0.10**	**−0.18**
	3.34	3.42	4.47	2.54	−0.30	−0.53
GAM-100 (10^{-8})	**0.75**	**0.87**	**1.08**	**0.65**	**−0.48**	**−0.46**
	2.64	3.06	3.66	2.31	−1.62	−1.60
NS-100	**0.71**	**0.78**	**0.98**	**0.65**	**−0.57**	**−0.52**
	1.98	2.18	2.61	1.83	-1.53	−1.45
PM$_{2.5}$						
GAM-30 (10^{-8})	**1.67**	**1.19**	**1.85**	**0.76**	**0.19**	**0.07**
	3.31	2.41	3.53	1.53	0.36	0.14
GAM-100 (10^{-8})	**1.38**	**0.75**	**1.14**	**0.62**	**~0**	**−0.39**
	3.13	1.71	2.41	1.41	0.002	−0.90
NS-100	**1.49**	**0.77**	**1.03**	**0.50**	**−0.19**	**−0.47**
	2.70	1.45	1.76	0.90	−0.33	−0.86

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

Table 18. Hospital Admissions for Chronic Obstructive Pulmonary Disease (COPD) in Los Angeles County: Estimated Percentage Change (log relative risk $\times$ 100) in Daily COPD Admissions Associated with Increases of 1 ppm CO and 10 ppb NO$_2$[a]

	Lag (days)					
	0	1	2	3	4	5
CO						
GAM-30 (10^{-8})	**5.48**	**5.67**	**5.90**	**5.28**	**4.59**	**4.10**
	17.67	18.22	19.01	16.94	14.50	12.80
GAM-100 (10^{-8})	**2.37**	**2.41**	**2.41**	**1.81**	**1.38**	**1.07**
	8.67	8.73	8.76	6.58	4.94	3.82
NS-100	**2.28**	**2.29**	**2.32**	**1.74**	**1.30**	**1.00**
	5.65	5.50	5.33	4.10	3.16	2.46
NO$_2$						
GAM-30 (10^{-8})	**2.84**	**2.74**	**2.74**	**2.08**	**1.57**	**1.22**
	13.32	12.71	12.60	9.57	7.21	5.62
GAM-100 (10^{-8})	**1.80**	**1.67**	**1.49**	**0.83**	**0.49**	**0.27**
	9.60	8.82	7.76	4.34	2.57	1.42
NS-100	**1.78**	**1.63**	**1.44**	**0.79**	**0.44**	**0.22**
	7.72	6.92	5.95	3.31	1.87	0.94

[a] For each model, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic. GAM-30(10^{-8}) = GAM with 30 *df* smoothing splines for temporal trends and the more stringent convergence criteria; GAM-100(10^{-8}) = GAM with 100 *df* smoothing splines and the more stringent convergence criteria; NS-100 = GLM with 100 *df* natural splines for temporal trends.

Table 19. Two-Pollutant Analyses (NO_2 and either PM_{10} or $PM_{2.5}$) of Hospital Admissions for Chronic Obstructive Pulmonary Disease (COPD) in Los Angeles County: GAM Using 100 *df* and More Stringent Convergence Criteria: Estimated Percentage Change (log relative risk $\times$ 100) in Daily COPD Admissions Associated with Increases of 10 µg/m³ PM_{10} and $PM_{2.5}$ and 10 ppb SO_2[a]

	Lag (days)			
	0	1	2	3
PM_{10} & NO_2				
PM_{10}	**−0.03**	**0.38**	**0.71**	**−0.05**
	−0.07	0.99	1.79	−0.12
NO_2	**1.72**	**1.14**	**0.91**	**1.64**
	3.18	2.11	1.63	3.00
$PM_{2.5}$ & NO_2				
$PM_{2.5}$	**0.42**	**−0.04**	**0.35**	**−1.09**
	0.62	−0.05	0.50	−1.65
NO_2	**1.51**	**1.24**	**1.31**	**2.86**
	2.07	1.71	1.70	4.03

[a] For each pollutant, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic.

increased smoothing, however, NO_2 remained significant at lag 0, became significant at lag 1, and lost significance at lag 5. In the original analyses I had reported no significant associations with SO_2. While this conclusion remained unchanged with the revised convergence criteria, with increased smoothing the coefficient for SO_2 became strongly significant ($t = 3.2$) at lag 0 (Table 20).

In my original paper I had not reported any two-pollutant analyses with PM_{10} and a gas. In the current analyses, when using models with 100 *df* smoothers for trends to analyze PM_{10} and SO_2 jointly, the coefficient of both pollutants were substantially attenuated and became insignificant at all lags. In analyses with NO_2, likewise, the coefficients for both pollutants were attenuated and insignificant. However, the coefficient for NO_2 remained nearly significant ($t = 1.91$) at lag 1 (results not shown).

DISCUSSION

In this report I have presented limited sensitivity analyses of data originally analyzed in a series of papers (Moolgavkar, 2000a,b,c). I began with the basic models I had developed in the previous analyses and reran them with the more stringent convergence criteria in GAM and also using natural splines in GLM. Some investigators at the US Environmental Protection Agency Workshop on GAM-Related Statistical Issues in PM Epidemiology (November 4–6, 2002, Research Triangle Park NC) had suggested that it might be necessary to develop the basic models again using the more stringent convergence criteria. I have not had the time to investigate this issue.

My results confirmed what others reported at the EPA workshop: changes in the convergence criteria and the use of GLM instead of GAM can, but do not always, substantially impact the results of the analyses and their interpretation. I summarize my conclusions as follows.

Table 20. Hospital Admissions for Chronic Obstructive Pulmonary Disease (COPD) in Cook County for GAM Using 100 *df* and More Stringent Convergence Criteria: Estimated Percentage Change (log relative risk $\times$ 100) in Daily COPD Admissions Associated with Increases of 10 µg/m³ PM_{10}, 1 ppm CO, and 10 ppb NO_2 and SO_2[a]

	Lag (days)					
	0	1	2	3	4	5
PM_{10}	**0.64**	**0.48**	**−0.06**	**−0.01**	**−0.05**	**−0.07**
	2.17	1.67	−0.21	−0.01	−0.16	−0.23
CO	**2.11**	**2.85**	**1.14**	**1.05**	**0.43**	**0.34**
	1.62	2.16	0.86	0.79	0.33	0.26
NO_2	**2.04**	**1.57**	**0.85**	**0.91**	**1.00**	**0.87**
	2.99	2.33	1.27	1.35	1.45	1.30
SO_2	**4.87**	**1.99**	**−0.11**	**−0.77**	**0.21**	**1.30**
	3.18	1.32	−0.07	−0.51	0.14	0.89

[a] For each pollutant, the upper row shows the estimated coefficient in bold type, the lower row the *t*-statistic.

1. In both Los Angeles and Cook Counties, the coefficients for individual pollutants and the associated *t*-statistics were substantially affected by making the convergence criteria more stringent in the GAM analyses. The general trend was a decrease in the estimated coefficient and the value of the *t*-statistic. This effect became more pronounced with increased smoothing of one of the covariates (temporal trends). These results are dramatically illustrated in Table 1, which shows the results of analyses of total nonaccidental mortality in Los Angeles County. I found similar, although not as striking, results in Cook County.

2. As judged by the AIC, models with 100 *df* for temporal trends generally, but not in all cases, described the data better than models with 30 *df* for temporal trends. The results with these two schemes for smoothing sometimes differed substantially. So the amount of smoothing of temporal trends can substantially affect the interpretation of analyses. GAMs fit the data better, as judged by the AIC, than GLMs employing natural splines with the same degree of freedom for temporal trends. However, GLMs yielded lower risks associated with pollutants and smaller *t*-statistics. As the tables show, these differences can be substantial.

3. I conducted limited sensitivity analyses in Los Angeles County using different amounts of smoothing of weather-related covariates (temperature, relative humidity). The coefficients and *t*-statistics appeared to be robust. This finding may not hold, however, in other metropolitan areas with more extremes of temperature and relative humidity or with greater seasonal variation in pollutant concentrations.

4. Given that different analytic strategies can substantially affect the estimates of effects of individual pollutants I believe that no numerical estimates are very meaningful. The patterns of association appear to be robust, however. For example, in Los Angeles (with the exception of COPD admissions with which NO_2 appeared to show the most robust association) it is clear that CO was the best single index of air pollution associations with health endpoints, far better than the mass concentration of either PM_{10} or of $PM_{2.5}$. In Cook County the results were not so clear cut. However, any one of the gases was at least as good an index of air pollution effects on human health as is PM_{10}.

REFERENCES*

Moolgavkar SH. 2000a. Air pollution and daily mortality in three US counties. Environ Health Perspect 108:777–784.

Moolgavkar SH. 2000b. Air pollution and hospital admissions for chronic obstructive pulmonary disease in three metropolitan areas in the United States. Inhalation Toxicol 12(Suppl 4):75–90.

Moolgavkar SH. 2000c. Air pollution and hospital admissions for diseases of the circulatory system in three US metropolitan areas. J Air Waste Manage Assoc 50:1199–1206.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
CO	carbon monoxide
COPD	chronic obstructive pulmonary disease
CVD	cardiovascular disease
df	degree of freedom
GAM	generalized additive model
GLM	generalized linear model
NO_2	nitrogen dioxide
PM	particulate matter
PM_{10}	particulate matter less than 10 µm in diameter
$PM_{2.5}$	particulate matter less than 2.5 µm in diameter
SO_2	sulfur dioxide

* Bold type identifies publications containing the original analyses revised in this short communication report.

Coarse Particles and Daily Mortality in Coachella Valley, California

Bart D Ostro, Rachel Broadwin, and Michael J Lipsett

ABSTRACT

Few studies have investigated associations of ambient coarse particles with daily mortality. In our original analysis using a generalized additive model (GAM*), we found associations between cardiovascular mortality and both particulate matter < 10 µm in median aerodynamic diameter (PM_{10}) and coarse particles (Ostro et al 2000). Here we examine the impacts of: (1) more stringent convergence criteria; (2) the use of a generalized linear model (GLM), which provides an unbiased estimate of the regression standard error; and (3) alternative lags, degrees of freedom (*df*) in the smoothers, and other model specifications. The tighter convergence criteria and the GLM generated regression estimates and standard errors similar to the original estimates, as did varying the degrees of freedom in the smoothers. Estimates of multiday exposures resulted in increases in the risk estimates associated with both PM_{10} and coarse particles. Therefore, this sensitivity analysis strongly supports the original findings of an association between mortality and a particle mix dominated by the coarse fraction.

INTRODUCTION

The Coachella Valley (*the Valley*), a popular resort and retirement destination, is a desert area with hot summers and mild winters. The Valley is roughly 100 miles east of Los Angeles and includes the cities of Palm Springs in the north and Indio in the south. It is surrounded by mountains on the north, east and west and bordered by the Salton Sea to the south. Coarse particles (PM_{10} minus $PM_{2.5}$ [particulate matter < 2.5 µm in median aerodynamic diameter]) comprise a significant percentage of the particulate mass;

approximately 65% of PM_{10} is larger than 2.5 µm in aerodynamic diameter, and daily variation in PM_{10} is driven largely by changes in the coarse fraction (Ostro et al 2000). Our original analysis (Ostro et al 2000) found an association between daily changes in cardiovascular mortality and both PM_{10} and coarse particles. Therefore, this reanalysis focused on cardiovascular mortality and explored the sensitivity of the results to: (1) corrected convergence criteria for GAM; (2) use of a generalized linear model (GLM); (3) time and weather smoothing with different degrees of freedom and (4) alternative lag structures.

DATA

In the original analysis of air pollution and mortality, we used data on daily mortality from the California Department of Health Services, Health Data and Statistics Branch, for the period January 1, 1989 through December 10, 1998 (Ostro et al 2000). Because the population of the Coachella Valley varies throughout the year, with a large influx of tourists during the winter and early spring, deaths of individuals listed with a permanent address in zip codes outside the Valley were excluded from the analyses. Daily counts of total deaths (minus accidents and homicides) were aggregated as well as total daily counts of deaths from respiratory and circulatory diseases (*International Classification of Diseases, Ninth Revision* [ICD-9] codes 460–519 and 393–440, respectively).

Pollutant data were obtained from the South Coast Air Quality Management District, which operates and maintains two fixed-site monitoring stations in the Valley. At the Indio monitoring site, ozone and PM_{10} were monitored continuously throughout the study period. Hourly PM_{10} data were collected using a β-attenuation monitor (model FH62I-N, Graseby-Andersen). The monitors in Palm Springs, situated approximately 25 miles to the northwest, measured ozone, nitrogen dioxide and carbon monoxide every hour, while PM_{10} was collected every sixth day. Supplementary data on $PM_{2.5}$ were collected from April 1996 through November 1998 at both sites using β-attenuation monitors, while an

---

additional continuous monitor for PM_{10} was deployed at the Palm Springs site during this latter interval.

Since $PM_{2.5}$ data were only available for 2.5 years, predictive models were estimated in the original analysis for both $PM_{2.5}$ and the coarse particle fraction (CF [PM_{10}–$PM_{2.5}$]) in an attempt to obtain measures of these pollutants for the full 10 years. Because the PM_{10} data were more complete at Indio ($n = 2990$) than at Palm Springs ($n = 1268$), we used the former for most of the subsequent analyses. Ultimately, the coarse fraction was estimated as a cubic function of PM_{10} ($R^2 = 0.95$) and generated for the full 10 years. For $PM_{2.5}$, predictive models could not be successfully estimated. Therefore, only the 2.5 years of measured values of $PM_{2.5}$ were used.

To adjust for possible effects of weather on mortality, we obtained daily meteorological data collected at the airport in Palm Springs from the National Climatic Data Center (Asheville NC). These data consisted of daily precipitation, minimum and maximum temperatures, dewpoint range and relative humidity.

METHODS

In the original paper, we used a locally weighted smoother (LOESS) of time to reduce any autocorrelation and seasonal patterns in the mortality time series. Visual inspection of the mortality data, however, suggested only a modest seasonal pattern, particularly for cardiovascular mortality. Therefore, only 10 *df* were originally employed in the smooth of time. However, in the original analysis, after examining regressions using from 10 to 60 *df*, we found that the overall results were relatively insensitive to the degrees of freedom used in the smoother. To further develop our regression model, we determined the best fit of several covariates prior to the entry of air pollution variables into the model. We examined, in turn, temperature, humidity and dewpoint (including lags of up to four days for each meteorological variable), and day of the week. We also examined whether any smoothers of these weather variables improved the fit, based on the Akaike information criterion (AIC).

Once the covariates most strongly associated with daily total mortality were determined, each pollutant was added separately to the model. Contemporaneous exposure and lags of up to 4 days, as well as a cumulative 4-day moving average, were examined to allow for a delayed effect of exposure. In this reanalysis, we used the same basic model to test for the impact of tighter convergence criteria and use of GLM. However, we undertook extensive sensitivity analyses to examine how alternative specifications and

degrees of freedom of the smoothed weather variables would affect the GLM results.

The usual application of GAM in air pollution time-series analyses has recently been reported to produce inaccurate results (HEI 2002; Domenici et al 2002). Specifically, two problems were discovered. First, the default convergence criteria for the GAM were too large, so that maximum likelihood estimates were not necessarily being obtained. Second, the estimated standard error in the GAM (even if corrected convergence criteria were used) could be too small, producing incorrect tests of significance (Ramsay et al 2003). Thus, as typically applied, GAM analyses could result in biased coefficients, which would in some cases be found to be spuriously significant.

To assess the potential impacts of GAM in our previous analyses, we conducted additional sensitivity analyses of our results. We focused on instances where positive associations between air pollution and mortality had been observed in order to reduce the likelihood of a type I error (incorrectly rejecting the null hypothesis of no effect of air pollution). The sensitivity analyses were fourfold, including: (1) application of tighter convergence criteria as suggested by the authors of GAM in S-Plus (Insightful Corp 2002); (2) use of a natural spline GLM, which provides correct standard errors of the estimated coefficients; (3) additional sensitivity analysis of the degrees of freedom used in LOESS; and (4) specification of alternative lag structures. The natural spline GLM is a parametric additive model that fits piecewise polynomials to segments of the data divided into separate regions by knots.

We did not attempt to reestimate the entire set of results originally presented. Instead, we examined whether the essential findings were affected by the reported problems with GAM. In order to more clearly observe any changes in the effect estimates, we present here the estimated β coefficients and *t*-statistics for the pollutant of interest. All of the analyses were conducted using S-Plus version 6.1.

RESULTS

Table 1 indicates the means and ranges of daily mortality, air pollution concentrations and meteorological variables. Mean daily mortality counts were 5.8 for all causes, 2.7 for cardiovascular causes, and 0.52 for respiratory causes. Based on the monitor in Indio, mean 24-hour concentrations of PM_{10}, coarse and fine particles (with different numbers of observations for each due to collection schedules and monitor malfunctions) were approximately 47, 26 and 17 $\mu g/m^3$, respectively. Gaseous pollutant concentrations

Table 1. Descriptive Statistics for Daily Mortality, Pollutant Concentrations, and Temperature, Coachella Valley, California: Jan 1, 1989 through Dec 10, 1998[a]

Variable	n (days)	Mean	Min	Max
Total mortality	3677	5.8	0	17
Cardiovascular mortality	3677	2.7	0	10
Respiratory mortality	3677	0.52	0	5
PM_{10} ($\mu g/m^3$, 24-hr mean, Indio)	3011	47.4	3	417
$PM_{2.5}$ ($\mu g/m^3$, 24-hr mean, Indio)	1041	16.8	5	48
Coarse particles ($\mu g/m^3$, 24-hr mean, Indio)	789	25.8	0	164
Predicted coarse particles ($\mu g/m^3$, 24-hr mean, Indio)	2990	30.5	0	418
Ozone (pphm, 1-hr max, Palm Springs)	3558	6.7	0	19
Nitrogen dioxide (pphm, 24-hr mean, Palm Springs)	3421	2.0	0	6
Carbon monoxide (ppm, 8-hr mean, Palm Springs)	3502	0.3	0	2.25
Max daily temperature (°F)	3462	88.7	49	125

[a] Min = minimum, max = maximum, pphm = parts per hundred million.

were low (carbon monoxide and nitrogen dioxide) to moderate (ozone).

The best de-trended model developed to explain variations in mortality prior to the inclusion of pollutant variables included maximum temperature (unlagged), indicator variables for Monday and Saturday, and a smooth of time with 10 *df*. In the original findings for cardiovascular mortality, associations were found for both PM_{10} and coarse particles, each lagged 0, 1, or 2 days (eg, for a 0-day lag: PM_{10} relative rate [RR] = 1.03, 95% confidence interval [CI] = 1.01–1.05 for an interquartile range [IQR] of 25 $\mu g/m^3$; coarse particles RR = 1.02, 95% CI = 1.01–1.04 for an IQR of 20 $\mu g/m^3$). No association with cardiovascular mortality was found for $PM_{2.5}$ or any of the gaseous pollutants. However, only 2.5 years of data were available for $PM_{2.5}$ versus 10 years for measured PM_{10} and estimated coarse particles.

Table 2 provides a summary of the results for both original estimates and estimates when the more stringent convergence criteria were applied. These criteria (using the GAM in S-Plus) included: `epsilon = 1`[-8], `maxit= 1000`, `bf.epsilon = 1`[-8], `bf.maxit = 1000`. Epsilon is the change in log likelihood in the local scoring and

Table 2. Effects of Alternative Convergence Criteria and Models on Coefficients Relating Particle Metrics to Daily Cardiovascular Mortality, Coachella Valley, California: Jan 1, 1989 through Dec 10, 1998 (β Coefficients and Standard Errors [SEs] × 1000)

Lag	PM_{10} β	SE	*t*-statistic	Coarse Particles β	SE	*t*-statistic	$PM_{2.5}$ β	SE	*t*-statistic
GAM Using Original Convergence Defaults for LOESS									
0	1.09	0.38	2.83	1.23	0.43	2.88	−5.60	3.38	−1.66
1	0.75	0.38	1.97	0.86	0.43	2.01	−1.42	3.26	−0.44
2	0.73	0.38	1.91	0.83	0.43	1.94	−2.29	3.22	−0.71
3	−0.17	0.40	−0.41	−0.14	0.46	−0.30	−0.98	3.19	−0.31
4	−0.32	0.41	−0.79	−0.40	0.47	−0.86	3.73	3.12	1.20
GAM Using New Convergence Criteria for LOESS[a]									
0	1.07	0.38	2.82	1.16	0.45	2.58	−5.80	3.40	−1.71
1	0.74	0.38	1.95	0.89	0.45	1.98	−1.30	3.30	−0.39
2	0.72	0.38	1.89	0.90	0.44	2.05	−2.30	3.20	−0.72
3	−0.19	0.40	−0.48	−0.25	0.48	−0.52	−0.87	3.20	−0.27
4	−0.33	0.41	−0.80	−0.37	0.48	−0.77	3.90	3.10	1.26
GLM Using Natural Spline									
0	0.99	0.38	2.57	1.08	0.46	2.36	−5.74	3.45	−1.66
1	0.69	0.39	1.77	0.82	0.45	1.81	−1.57	3.36	−0.47
2	0.67	0.39	1.73	0.84	0.45	1.88	−2.40	3.30	−0.74

[a] `Epsilon = 1e-8, maxit= 1000, bf.epsilon = 1e-8, bf.maxit = 1000`; regression model includes unlagged maximum temperature, indicator variables for Monday and Saturday, and a LOESS of time (*df* = 10).

backfitting models and maxit is the maximum iterations in the two models. Comparing the original with the newer estimates using the tighter convergence criteria showed little difference in the estimated pollution coefficients or their standard errors in any of the single-day lags. As in the original analysis, associations were detected for 0-day, 1-day, and 2-day lags for both PM_{10} and coarse particles, but not for $PM_{2.5}$. The coefficients for PM_{10} and coarse particles were slightly lower with the revised convergence criteria. These results were confirmed when even more stringent convergence criteria were applied (eg, `epsilon` and `bf.epsilon = 1`$^{-10}$) (data not shown). We next tested the sensitivity of the results to use of a natural spline in GLM (Table 2). Again, little change in the estimates or standard errors was observed.

Table 3 summarizes the results for coarse particles (2-day lag) using time and temperature smoothers with varying degrees of freedom in both GAM and GLM, and using multipollutant models. Models 1, 2 and 5 are reported in Table 2 but are repeated in Table 3 for comparison. Models 3 and 4 both increased the degrees of freedom in the smooth of time, while model 4 also added a smooth

of temperature in GAM. Models 6 and 7 were similar to models 3 and 4, using GLM instead of GAM, while model 8 used the B-spline option in GLM. The B-spline is similar to the natural spline parametric model except that the function need not be linear at the boundaries of the data. Models 9 and 10 were two-pollutant models, incorporating ozone or nitrogen dioxide. In general, these regressions demonstrated slight to modest increases in the magnitude and statistical significance of the estimated coefficients for coarse particles in relation to the original GAM results.

Table 4 summarizes the results for coarse particles using alternative lags after eliminating days with extreme values for wind speed and maximum daily temperature. Windy days are often associated with high coarse particle concentrations and may also be days when people significantly reduce their time outdoors. Cumulative averages of 3 days (the average of lags 0 through 2) and 5 days (the average of lags 0 through 4) generated estimates markedly greater than the estimate based on a single 2-day exposure lag. Deleting 5% of days with the highest 1-hour maximum wind speed or 1-hour maximum temperature also increased the estimated effect of coarse particles.

Table 3. Effects of Alternative Model Specifications on Coefficients Relating Coarse Particle Concentrations to Daily Cardiovascular Mortality, Coachella Valley, California: Jan 1, 1989 through Dec 10, 1998 (2-day Lag, β Coefficients and Standard Errors × 1000)

GAM	
Original	0.83 (0.43)
With new criteria	0.90 (0.44)
With new criteria and smooth (day, $df = 60$)	0.92 (0.44)
With 2 smooths (day, $df = 60$; temp, $df = 20$)	0.97 (0.45)
GLM	
Similar specification as original GAM	0.84 (0.45)
Natural spline with smooth (day, $df = 60$)	0.88 (0.47)
Natural spline with 2 new smoothers (day, $df = 60$; temp, $df = 20$)	0.93 (0.48)
B-spline with new smooth (day, $df = 60$)	0.94 (0.47)
GAM with new criteria plus ozone	1.00 (0.45)
GAM with new criteria plus NO_2	1.05 (0.47)

Note: Basic regression model includes unlagged maximum temperature, indicator variables for Monday and Saturday, and a LOESS of time ($df = 10$).

Table 4. Effects of Alternative Lag Structures and Deleting Days with Extreme Values on Coefficients Relating Coarse Particle Concentrations to Daily Cardiovascular Mortality, Coachella Valley, California: Jan 1, 1989 through Dec 10, 1998 (2-Day Lag, Coefficients and Standard Errors × 1000)

Original GAM, lag2	0.83 (0.43)
GAM with new convergence criteria, lag2	0.90 (0.44)
GAM with new convergence criteria, lag0	1.16 (0.45)
GAM with new convergence criteria, lag02	2.11 (0.65)
GAM with new convergence criteria, lag04	1.59 (0.81)
GAM with new convergence criteria, lag0 (−5% max wind speed)	1.29 (0.55)
GAM with new convergence criteria, lag0 (−5% max temp)	1.28 (0.47)

Note: Basic regression model includes unlagged maximum temperature, indicator variables for Monday and Saturday, and a LOESS of time ($df = 10$). Lag0 = single-day unlagged exposure; Lag2 = two-day lagged exposure; Lag02 = cumulative average of 0-, 1-, and 2-day lagged exposures; Lag04 = cumulative average of 0-, 1-, 2-, 3- and 4-day lagged exposures; (−5% max wind) = model without 5% of the days with the highest maximum wind speed; (−5% max temp) = model without 5% of the days with the highest maximum temperature.

In additional sensitivity analyses, we found no differences from the original results for all-cause or respiratory mortality or for other pollutants or lags (data not shown).

DISCUSSION

PM_{10} in the Coachella Valley is dominated by the coarse fraction. In 2.5 years of data collected in Palm Springs and Indio, the coarse fraction comprised approximately 60% of PM_{10} mass (Ostro et al 1999). Daily correlations between PM_{10} and the coarse fraction were high in both areas: $r = 0.94$ in Palm Springs and 0.97 in Indio, while the correlations between PM_{10} and $PM_{2.5}$ were 0.68 and 0.46, respectively. Thus, daily variation in PM_{10} during the 10-year study period was likely driven primarily by changes in the coarse fraction, with fine particles exerting a lesser influence. The magnitude of the estimated relative risks of cardiovascular mortality associated with PM_{10} and coarse particles is about 1% per 10 μg/m^3 PM_{10}.

Our reanalysis using tighter convergence criteria for GAM, as well as GLM in S-Plus, indicated that the original results (Ostro et al 2000) were essentially unchanged. Further, varying degrees of freedom for smoothers of time or temperature had little effect on the original estimates and, if anything tended to increase the estimates of effect slightly. Our results differ markedly from those of Schwartz et al (1996), who reported associations with fine, but not coarse, particles on mortality in 6 cities. One possibility for the discrepancy with their results is the relatively high concentration of the coarse fraction in our sample—a 2.5-year mean of 26 μg/m^3. In the Schwartz et al (1996) analysis of 6 cities, associations between coarse particles and mortality were observed only for Steubenville, the city with the highest long-term average concentration of coarse particles (16 μg/m^3). Most cities that have been examined in daily time-series mortality studies have relatively low concentrations of coarse particles, often around 10 μg/m^3 or less. An exception to this is the study by Mar et al (2000) in Phoenix, Arizona, where the mean coarse particle concentration was reported to be 34.5 μg/m^3. In that locale, coarse particles were also associated with cardiovascular mortality.

A second possible explanation of the association of mortality with coarse particles is that the warm climate during most of the year, including the winter, might encourage residents to keep windows and doors open, thereby facilitating penetration of coarse particles into the indoor environment and increasing total exposure. Third, the large pool of retirees in the Coachella Valley area may constitute a population particularly susceptible to particulate matter

air pollution, including coarse particles. Fourth, there may be a large fraction of the coarse particle mass in an inter-modal size range (1.5–4 μm), which may carry pathophysiological consequences similar to fine particles in other settings. Finally, coarse particles may contain silica or other reactive metals or biological contaminants such as endotoxin that might initiate inflammatory processes thought to be associated with cardiovascular events (Air Resources Board 2002).

In summary, our reanalysis, using convergence criteria and model specifications different from our original GAM-based estimates, supports the original findings of an association between coarse particles and daily cardiovascular mortality in this desert setting.

REFERENCES*

Air Resources Board. 2002. Staff Report: Public Hearing to Consider Amendments to the Ambient Air Quality Standards for Particulate Matter and Sulfates. Prepared by the Staff of the Air Resources Board and the Office of Environmental Health Hazard Assessment. May 3, 2002. (www.arb.ca.gov/research/aaqs/std-RS/pm-final/pm-final.htm.)

Dominici F, McDermott A, Zeger SL Samet JM. 2002. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiol 156:193–203.

Health Effects Institute. 2002. Letter from Dan Greenbaum, May 30, 2002. (www.healtheffects.org/Pubs/NMMAPSletter.pdf)

Insightful Corporation. 2002. How to tighten default convergence criteria for GAM and GLM. Insightful Corporation, Seattle WA. (www.insightful.com/support/faqdetail.asp?FAQID=139&IsArchive=0).

---

* Bold type identifies publication containing the original analyses revised in this short communication report.

Mar TF, Norris GA, Koenig JQ, Larson TV. 2000. Associations between air pollution and mortality in Phoenix, 1995–1997. Environ Health Perspect 108: 347–353.

Ostro BD, Broadwin R, Lipsett MJ. 2000. Coarse and fine particles and daily mortality in the Coachella Valley, California: A follow-up study. J Exposure Anal Environ Epidemiol 10:412–419.

Ostro BD, Hurley S, Lipsett MJ. 1999. Air pollution and daily mortality in the Coachella Valley CA: A study of PM_{10} dominated by coarse particles. Environ Res 81:231–238.

Ramsay T, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Schwartz J, Dockery, DW, Neas LM. 1996. Is daily mortality associated specifically with fine particles? J Air Waste Manage Assoc. 46: 927–939.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
CI	confidence interval
df	degrees of freedom
GLM	generalized linear model
IQR	interquartile range
LOESS	locally weighted smoother
PM_{10}	particulate matter < 10 µm in median aerodynamic diameter
$PM_{2.5}$	particulate matter < 2.5 µm in median aerodynamic diameter
r	bivariate correlation coefficient
R^2	multivariate coefficient of determination
RR	relative rate

Sensitivity Analyses of Regional Differences in Short-Term Effects of Air Pollution on Daily Mortality in APHEA Cities

E Samoli, J Schwartz, A Analitis, Y Petasakis, B Wojtyniak, G Touloumi, C Spix, F Balducci, S Medina, G Rossi, J Sunyer, HR Anderson, and K Katsouyanni

ABSTRACT

Short-term effects of air pollution on daily mortality in 8 western and 5 central-eastern European cities have been reported previously as part of the Air Pollution and Health: A European Approach (APHEA*) 1 project. One intriguing finding was that the effects were lower in central-eastern European cities. We had published results from a sensitivity analysis comparing the original methods, which used sinusoidal terms for seasonal control and polynomial terms for meteorologic variables, and the use of generalized additive models (GAMs) (Samoli et al 2001).

Recently there has been some concern about the use of the S-Plus GAM function. It has been shown that its default convergence criteria are not strict enough (resulting in biased estimates of regression coefficients) and that the function underestimates the variances of the parameters.

In this short communication report, prepared in response to a request by the US Environmental Protection Agency (EPA), we present results of the revised analysis of the data published in the sensitivity paper (Samoli et al 2001). This analysis used the S-Plus GAM with more stringent convergence criteria and Poisson regression with natural splines to control confounding.

When the GAM with more stringent convergence criteria was applied, the pooled estimates for the analyzed pollutants were reduced by less than 10%. The results from the analyses using natural splines were consistently smaller than the results obtained by GAM with more stringent convergence criteria.

INTRODUCTION

Short-term effects of air pollution on daily deaths have been investigated in a large European multicenter study, the APHEA project. The first part of the project included data from 15 cities. Of these, 5 were central-eastern European cities for which exposure data on sulfur dioxide (SO_2) and black smoke levels were available. In the original APHEA project, data from each city were analyzed according to a standardized protocol (Katsouyanni et al 1996) with sinusoidal terms for seasonal control and polynomial terms for meteorologic variables. One intriguing finding was that the effects were lower (although still statistically significant in most instances) in central-eastern European cities (Katsouyanni et al 1997, Zmirou et al 1998) compared with western European cities. A possible explanation for this finding is that the (parametric) model for seasonal control may have fit the eastern-cities data less well because of a higher, more variable rate of respiratory illness or because of differences in air pollution concentrations.

Owing to the observed heterogeneity in the findings, the APHEA group had undertaken a sensitivity analysis of the original results using GAM to test the adequacy of seasonal control and to provide a basis for comparison with the results of the new APHEA2 project.

Recently, Dominici and colleagues (2002) identified a problem with the default convergence criteria of GAM in S-Plus software. Independently, Ramsay and associates (2003) found that this GAM underestimated parameter variances. In response to these findings and a request by the EPA, we reanalyzed our data using the S-Plus GAM with more stringent convergence criteria and Poisson regression with natural splines to control seasonality and meteorologic variables.

DATA AND METHODS

The data used in this analysis concern 7 western European cities (Athens, Barcelona, Cologne, London, Lyon, Milan, and Paris) and 5 central-eastern European cities (Bratislava, Cracow, Lodz, Poznan, and Wroclaw). The pollutants studied were SO_2 (24 hour) (available for all cities) and black smoke (available for 4 western and 4 central-eastern cities). The daily number of deaths from all causes except external causes was the health outcome. The data covered at least 5 consecutive years from 1980 to 1992 for each city. Details about these data have been published elsewhere (Katsouyanni et al 1997, Zmirou et al 1998). The analyses were restricted to days when SO_2 and black smoke levels were less than 200 μg/m^3, for which roughly linear associations with the logarithm of expected mortality are observed.

METHODS

A hierarchical modeling approach was used. In the first stage, city-specific models were built; in the second, fixed- and random-effects metaanalyses were used with inverse variance weighting. In the original APHEA1 individual-city models, sinusoidal terms were used to control seasonal patterns and polynomial terms for meteorologic variables (Katsouyanni et al 1996). In the published sensitivity analysis (Samoli et al 2001), GAM Poisson regression using locally weighted smoothers (LOESSs) were applied to control for seasonal patterns and meteorologic variables. S-Plus GAM function with default convergence criteria was used.

More specifically, in the GAM applied in the original sensitivity analysis, we used a LOESS for meteorologic variables with a smoothing window that minimized the Akaike information criterion (AIC). For long-term trends a window between 80 and 200 was chosen a priori. Within that range, for each city we chose the span that minimized the partial autocorrelation of the residuals (see Appendix Table A.1). Day of the week effects, holidays, and epidemics were controlled for using indicator variables. Robust regression was used to reduce the effect of any extreme observations on the regression results. In APHEA1 the lag of the pollutant that best fit the data was chosen. To maintain comparability, we used the same lags in the revised sensitivity analysis.

Once models were fit for each city, we summarized the results over all cities using inverse variance weighting. We computed separate summaries for eastern Europe, western Europe, and all cities. We examined heterogeneity by computing χ^2 statistics.

For the present reanalysis we applied the first-stage models used in our nonparametric analyses, changing the convergence criteria to more stringent ones. More specifically, we set the maximum number of iterations to 1000 and the difference of two successive coefficients to 10^{-14}. As requested by the EPA, we also used natural splines with the same degree of freedom as those used in the original GAM (with LOESS) to adjust for seasonal patterns and meteorologic variables.

RESULTS

Table 1 presents the individual city and pooled β coefficients and their standard errors for black smoke concentration using the original parametric method (sinusoidal control), the original GAM with the default convergence criteria, GAM with the more stringent convergence criteria, and natural splines (with the same degree of freedom as for the original nonparametric model). The results differed only slightly between the GAMs with more stringent and default convergence criteria (Table 2): the reduction in the pooled estimates was less than 10%, whereas the change in individual city estimates ranged from -12.8% to -3.5%. Compared with the original parametric method, the GAMs gave substantially higher fixed-effects pooled estimates for eastern-European cities but similar estimates for western-European cities (Table 2). The natural splines estimates differed widely from those of the original parametric model, higher for 4 cities and smaller for the other 4. Natural spline estimates were consistently smaller (by 30–60%) than those from the GAM with more stringent convergence criteria. The fixed effects pooled estimate was reduced by 47.5% (Table 2).

Standard errors of the coefficients from city-specific models generally increased for the original parametric model or natural splines compared with the GAMs. However, standard errors of the pooled estimates were quite similar across all models.

The results for SO_2 effects were broadly similar to those for black smoke effects (Tables 3 and 4).

Table 1. Black Smoke Effects on Daily Total Number of Deaths: Results from Alternative Models[a]

City	Sin/Cos β	Sin/Cos SE	Default β	Default SE	Strict β	Strict SE	ns[b] β	ns[b] SE
Athens	0.00048	0.00011	0.00086	0.00012	0.00079	0.00012	0.00045	0.00014
Barcelona	0.00068	0.00020	0.00054	0.00016	0.00049	0.00015	0.00034	0.00018
Bratislava	-	-	-	-	-	-	-	-
Cracow	0.00014	0.00008	0.00059	0.00011	0.00057	0.00011	0.00023	0.00013
Koln	-	-	-	-	-	-	-	-
Lodz	0.00013	0.00008	0.00047	0.00008	0.00042	0.00008	0.00022	0.00010
London	0.00112	0.00030	0.00073	0.00025	0.00068	0.00024	0.00041	0.00027
Lyon	-	-	-	-	-	-	-	-
Milano	-	-	-	-	-	-	-	-
Paris	0.00053	0.00015	0.00034	0.00012	0.00030	0.00011	0.00013	0.00014
Poznan	0.00003	0.00020	0.00024	0.00015	0.00022	0.00015	0.00015	0.00017
Wroclaw	0.00010	0.00013	0.00031	0.00014	0.00030	0.00014	0.00016	0.00015
Fixed-effects pooled estimate	0.00026	0.00004	0.00050	0.00004	0.00046	0.00004	0.00024	0.00005
Western Europe	0.00057	0.00008	0.00060	0.00007	0.00054	0.00007	0.00032	0.00008
Eastern Europe	0.00012	0.00005	0.00045	0.00006	0.00041	0.00006	0.00020	0.00007

[a] Sin/Cos = original parametric method; Default = original GAM using S-Plus with default convergence criteria; Strict = original GAM using S-Plus with more stringent convergence criteria; ns = model with natural splines.

[b] Same degree of freedom was used as for the original nonparametric model.

Table 2. Percent Difference in Black Smoke Regression Coefficients Between Alternative Models[a]

City	Strict vs Sin/Cos	Strict vs Default	ns vs Strict	ns vs Sin/Cos
Athens	65.4	−7.7	−43.3	−6.3
Barcelona	−27.3	−8.5	−31.4	−50.2
Cracow	309.2	−3.5	−59.1	67.2
Lodz	222.5	−10.8	−48.1	67.3
London	−39.3	−6.9	−40.0	−63.6
Paris	−44.1	−12.8	−55.1	−74.9
Poznan	640.2	−6.5	−32.5	399.4
Wroclaw	196.0	−3.8	−46.8	57.4
Fixed-effects pooled estimate	78.5	−7.4	−47.5	−6.3
Western Europe	−4.6	−9.2	−41.9	−44.6
Eastern Europe	243.8	−7.3	−51.8	65.6

[a] Strict = original GAM using S-Plus with more stringent convergence criteria; Sin/Cos = original parametric method; Default = original GAM using S-Plus with default convergence criteria; ns = model with natural splines.

DISCUSSION

The GAM applied in this sensitivity analysis generally led to increases in the estimated pooled effects of SO_2 and black smoke on total mortality compared with the parametric methods (with either sinusoidal or natural splines to control for confounding). Although the changes in estimates for the central-eastern European cities were proportionally higher than those in the western-European cities, those for the central-eastern European cities remained lower by about 50% for both pollutants. The natural splines model results were closer to those from the parametric models with sinusoidal control: for SO_2, the natural splines model led to a small decrease in pooled estimates, whereas the effect of black smoke increased for central-eastern European cites and decreased for western-European cities. Since the models differed mainly in how they controlled confounding, these results highlight the importance of appropriate control for long-term trends and seasonality as well as for meteorologic variables. This issue needs to be investigated further.

The Samoli and colleagues (2001) paper concluded that part of the heterogeneity in air-pollution estimates between central-eastern and western European cities was due to inadequate control of seasonality but that the remaining heterogeneity should be investigated further. This conclusion remains after considering the revised analysis.

Table 3. SO_2 Effects on Daily Total Number of Deaths: Results from Alternative Models[a]

City	Sin/Cos		Default		Strict		ns[b]	
	β	SE	β	SE	β	SE	β	SE
Athens	0.00110	0.00016	0.00157	0.00014	0.00147	0.00014	0.00088	0.00019
Barcelona	0.00120	0.00027	0.00112	0.00022	0.00107	0.00022	0.00093	0.00024
Bratislava	−0.00051	0.00063	−0.00066	0.00047	−0.00069	0.00048	−0.00099	0.00057
Cracow	0.00031	0.00011	0.00048	0.00011	0.00046	0.00011	0.00021	0.00013
Koln	0.00028	0.00011	0.00035	0.00012	0.00033	0.00012	0.00023	0.00013
Lodz	0.00011	0.00014	0.00056	0.00010	0.00050	0.00010	0.00026	0.00012
London	0.00034	0.00017	0.00044	0.00015	0.00043	0.00015	0.00037	0.00018
Lyon	0.00120	0.00033	0.00190	0.00031	0.00188	0.00031	0.00168	0.00034
Milano	0.00041	0.00017	0.00077	0.00010	0.00065	0.00010	0.00028	0.00014
Paris	0.00053	0.00015	0.00040	0.00011	0.00038	0.00011	0.00031	0.00014
Poznan	0.00022	0.00014	0.00041	0.00014	0.00038	0.00014	0.00015	0.00016
Wroclaw	−0.00009	0.00017	0.00014	0.00017	0.00011	0.00017	−0.00018	0.00019
Fixed-effects pooled estimate	0.00038	0.00005	0.00061	0.00004	0.00055	0.00004	0.00031	0.00005
Western Europe	0.00057	0.00006	0.00073	0.00005	0.00067	0.00005	0.00051	0.00005
Eastern Europe	0.00017	0.00007	0.00043	0.00006	0.00040	0.00006	0.00014	0.00007

[a] Sin/Cos = original parametric method; Default = original GAM using S-Plus with default convergence criteria; Strict = original GAM using S-Plus with more stringent convergence criteria; ns = model with natural splines.

[b] Same degree of freedom was used as for the original nonparametric model.

REFERENCES*

Dominici F, McDermott A, Zeger S, Samet J. 2002. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiol 156: 193–203.

Katsouyanni K, Schwartz J, Spix C, Touloumi G, Zmirou D, Zanobetti A, Wojtyniak B, Vonk JM, Tobias A, Ponka A, Medina S, Bacharova L, Anderson HR. 1996. Short-term effects of air pollution on health: A European approach using epidemiologic time series data: the APHEA protocol. J Epidemiol Comm Health 50(suppl 1): S12–S18.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak B, Sunyer J, Bacharova L, Schouten JP, Ponka A, Anderson HR. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: Results from time series data from the APHEA project. BMJ 314: 1658–1663.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Samoli E, Schwartz J, Wojtyniak B, Touloumi G, Spix C, Balducci F, Medina S, Rossi G, Sunyer J, Bacharova L, Anderson HR, Katsouyanni K. 2001. Investigating regional differences in short-term effects of air pollution on daily mortality in the APHEA project: A sensitivity analysis for controlling long-term trends and seasonality. Environ Health Perspect 109:349–353.

Zmirou D, Schwartz J, Saez M, Zanobetti A, Wojtyniak B, Touloumi G, Spix C, Ponce de Leon A, Moullec Y, Bacharova L, Schouten J, Ponka A, Katsouyanni K. 1998. Time-series analysis of air pollution and cause specific mortality: A quantitative summary in Europe (APHEA study). Epidemiology 9: 495–503.

* Bold type identifies publication containing the original analyses revised in this short communicaiton report.

Table 4. Percent Difference in SO$_2$ Regression Coefficients Between Alternative Models[a]

	% Difference in β Coefficients			
	Strict vs Sin/Cos	Strict vs Default	ns vs Strict	ns vs Sin/Cos
City				
Athens	33.3	−6.6	−40.0	−20.0
Barcelona	−10.7	−4.3	−12.8	−22.1
Bratislava	34.3	3.8	45.1	94.8
Cracow	49.5	−2.7	−55.7	−33.8
Koln	16.9	−7.8	−28.3	−16.2
Lodz	358.2	−9.7	−49.4	132.1
London	27.2	−1.7	−13.6	9.9
Lyon	56.9	−0.9	−10.7	40.1
Milano	58.2	−15.3	−57.1	−32.1
Paris	−28.0	−4.6	−19.1	−41.7
Poznan	70.7	−7.3	−59.0	−30.1
Wroclaw	−217.3	−23.0	−269.0	95.8
Fixed-effects pooled estimate	46.0	−8.6	−44.0	−18.3
Western Europe	17.0	−8.8	−24.1	−11.2
Eastern Europe	133.7	−8.4	−63.9	−15.5

[a] Strict = original GAM using S-Plus with more stringent convergence criteria; Sin/Cos = original parametric method; Default = original GAM using S-Plus with default convergence criteria; ns = model with natural splines.

APPENDIX A

Table A.1. Variables and Degree of Freedom Included in City-Specific Models

City	Degree of Freedom			
	Years	Trend	Temperature	Humidity
Black Smoke				
Athens	5	23	9	5
Barcelona	7	34	3	2
Cracow	13	37	6	linear
Lodz	14	50	6	linear
London	5	28	4	linear
Paris	6	26	3	linear
Poznan	8	27	5	3
Wroclaw	11	34	4	linear
SO$_2$				
Athens	5	23	9	5
Barcelona	7	26	2	2
Bratislava	5	12	2	3
Cracow	13	36	6	linear
Koln	11	30	7	-
Lodz	14	50	6	linear
London	5	28	4	linear
Lyon	6	14	2	linear
Milano	10	57	9	2
Paris	6	26	3	linear
Poznan	8	27	5	3
Wroclaw	11	34	4	linear

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
APHEA	Air Pollution and Health: A European Approach
EPA	Environmental Protection Agency (US)
GAM	generalized additive model
LOESS	locally weighted smoother
SO$_2$	sulfur dioxide

Airborne Particles and Daily Deaths in 10 US Cities

Joel Schwartz

ABSTRACT

Generalized additive models (GAMs*) are commonly used to control for confounding in studies of the association of air pollution with counts of adverse events. Recently, it was pointed out that the default convergence criteria in most GAM software have been lax and that standard errors were not estimated correctly. Alternative approaches, including natural splines and penalized splines provide proper standard errors.

I have previously published a number of studies of the relation between PM_{10} (particulate matter < 10 μm in median aerodynamic diameter) and daily deaths in 10 US cities that had daily monitoring for PM_{10}. These included studies of the effects of PM_{10} averaged over the two days leading to death (including studies of effect modification and confounding) as well as studies of the distributed lag between PM_{10} and both all cause and cause-specific mortality. In this study I reanalyzed those associations using either GAM software with a stringent convergence criterion, natural splines, or penalized splines to control for season and weather variables.

INTRODUCTION

GAMs (Hastie and Tibshirani 1990) have been applied in many time-series studies of air pollution and mortality or morbidity. The wide use of GAMs in air pollution epidemiology is due to its flexibility in modeling nonlinear factors such as season and weather (Schwartz 1993 and 1994, Katsouyanni et al 2001, Daniels et al 2000).

The GAM function is available in the S-Plus statistical software (Mathsoft Inc, Seattle WA). Estimation in GAM (as

written for most current statistical packages) is based on a combination of the local scoring algorithm (to fit the non-Gaussian nature of the data) and the backfitting algorithm.

When the smoothing functions in the linear predictor are parametric, then the additive regression model is fit by using weighted least square and the GAM is equivalent to the generalized linear model (GLM). The backfitting algorithm is used within the local scoring algorithm iteration when several nonparametric smoothing functions are used in the model. With only one smooth function, backfitting is not required.

A recent report from Dominici and coworkers (Dominici et al 2002) indicated that the default convergence criteria used in the S-Plus function GAM has been relatively lax and may not guarantee convergence. This is not a software problem, merely a reminder that investigators need to pay attention to defaults in statistical software because they are not always appropriate for each problem. A separate paper (Ramsay et al 2003) reported that the S-Plus GAM function used a shortcut in estimating the covariance of the estimated coefficients that does not properly take into account the correlation between the exposure variables of interest and the smoothed functions of covariates. This error may result in biased estimates of the standard errors for the pollution variables.

These two findings have raised questions about the specific results of air pollution time series previously reported, and the reliability of the general approach and findings in particular. To address this, I have reanalyzed the data from a series of papers that evaluated different aspects of the association between PM_{10} and daily deaths in 10 US cities.

I selected ten US cities with roughly daily PM_{10} monitoring to provide a reasonable number of locations for a combined analysis. The cities were New Haven, Pittsburgh, Birmingham, Detroit, Canton, Chicago, Minneapolis–St Paul, Colorado Springs, Spokane, and Seattle. Daily deaths in the metropolitan county containing each city were extracted from National Center for Health Statistics mortality tapes for the years 1986 through 1993. I also computed separate daily counts of

---

---

deaths inside a hospital and deaths outside of a hospital. Minneapolis and St Paul were combined and treated as one city. Daily weather data were obtained for the same years, from the nearest airport weather station, and daily concentrations of PM_{10}, SO_2, O_3, and CO were obtained for those years from the US Environmental Protection Agency's Aerometric Information Retrieval System (AIRS) monitoring network. NO_2 was not available in enough cities to allow examination of that variable.

These papers addressed several aspects of the association between PM_{10} and daily deaths from nonexternal causes. The first paper (Schwartz 2000a) examined the association between PM_{10} and daily deaths in a large multicity study, looked at differences by season, and used a meta-regression technique to explore confounding by gaseous pollutants. Because it used cities with daily monitoring for PM_{10}, this study (unlike the NMMAPS 90-cities mortality study) was able to use a two-day average of PM_{10} to examine the association. The second paper (Schwartz 2000b) examined the distributed lag of the association between PM_{10} and daily deaths in the same cities. It demonstrated that effects of PM_{10} persist for one to two days after the day of exposure and that use of a single day's exposure, as in the NMMAPS mortality study, would be expected to underestimate the effect size. This mechanism likely accounts for the somewhat smaller estimates in the NMMAPS mortality study compared to other studies, almost all of which have used two or three day averages of exposure. Selection bias in the choice of cities has been hypothesized as an alternative explanation, but this explanation is unlikely because the APHEA2 study (Katsouyanni et al 2001), which also had no selection bias in the choice of cities, also reported larger effect sizes.

The final paper reanalyzed here (Braga et al 2001) examined how the time course of the association between PM_{10} and daily deaths varied by cause of death. The association was predominantly with the same day's exposure for deaths from myocardial infarction, with the exposures on the same or previous day for deaths from cardiovascular disease, and with the exposures on the two prior days for respiratory deaths. The effects dropped to zero for greater lags.

In this paper we report the reanalysis of these previous studies using GAM with locally weighted smoothers (LOESS) with a stringent convergence criterion, and either natural spline and penalized spline models, both of which provide correct standard errors.

Natural splines have been applied as an alternative to nonparametric smoothing (Dominici et al 2002), but these models are sensitive to knot locations and knots number. Penalized spline represents an attractive alternative that maintains the flexibility of nonparametric smoothing

without the problems associated with GAM. Penalized splines also are advantageous for model building because of its lack of sensitivity to knot locations. In addition, penalized spline software does not use backfitting.

METHODS

All three papers adopted a multilevel modeling approach. First separate regressions were fit in each city, and those results were then combined across cities in a second stage analysis that included different aspects in the different papers.

CITY-SPECIFIC ANALYSIS

For each city, the original analysis fit a generalized additive Poisson regression (Schwartz 1993) modeling the logarithm of the expected value (E) of daily deaths as a sum of smooth functions of the predictor variables.

It assumed that:

$$\log(E(Y)) = \beta_0 + S_1(X_1) + \ldots + S_p(X_p)$$

where Y is the daily count of deaths, $E(Y)$ is the expected value of that count, the X_i are the covariates and the S_i are the smooth (ie, continuously differentiable) functions. This analysis was done using daily deaths as the outcome in Schwartz (2000a, 2000b) and using cause-specific mortality in Braga et al (2001). In all cases the results were combined across the 10 cities as described below. The methods for choosing the smoothing parameters are described in the original papers. Covariates controlled for were temperature, dew point (or, depending on the paper, relative humidity), barometric pressure, season, and day of the week.

SECOND LEVEL ANALYSIS

All of the papers combined evidence across cities using meta-analytic techniques. These included simple inverse variance weighted averages and meta-regression with inverse variance weights. Fixed effect metaanalysis and metaregression were used in the original papers. One particular innovation of the Schwartz (2000a) paper was the use of staged regression analysis to control for confounding by copollutants in the second stage of the hierarchical model. This is described in more detail in the original paper. Briefly, consider first a Gaussian outcome:

$$Y_t = \beta_0 + \beta_1 X_t + \text{error}$$
(equation 1 in Schwartz [2000a])

where X_t is correlated with another pollutant Z_t. Therefore we may write

$X_t = \gamma_0 + \gamma_1 Z_t + \text{error}$
(equation 2 in Schwartz[2000a]).

What happens if we use Z_t as the exposure variable instead of X_t? Substituting (2) into (1), we have

$Y_t = \beta_0 + \beta_1\gamma_0 + \beta_1\gamma_1 Z_t + \text{error}$
(equation 3 in Schwartz [2000a]).

We have confounding by the omitted covariate X_t, and the coefficient of Z_t that we find will be proportional to γ_i, the slope of the association between X_t and Z_t. This provides a mechanism for estimating the effect of X, and by reversal, the effect of Z. We have generalized and shown that in the presence of measurement error in the exposures, this method has important advantages over the traditional two pollutant approach. For a more detailed demonstration of how this method is more resistant to measurement error than a two-pollutant model, see Schwartz and Coull (in press).

DISTRIBUTED LAG MODEL

Two of the papers (Schwartz 2000b, Braga et al 2001) used distributed lag models to assess the association between air pollution and death over several subsequent days. The motivation for the distributed lag model is the realization that air pollution can affect not merely deaths occurring on the same day, but on several subsequent days. The converse is therefore also true: Deaths today will depend on the same day effect of today's pollution levels and the one day lag effects of yesterday's PM_{10} concentrations, etc. Therefore, suppressing covariates and assuming Gaussian data for the moment, the unconstrained distributed lag model assumes:

$Y_t = \alpha + \beta_0 X_t + \ldots + \beta_q X_{t-q} + \varepsilon_t$
(equation 1 in Schwartz [2000b])

where X_{t-q} is the PM_{10} concentration q days before the deaths. The overall effect of a unit increase in air pollution on a single day is its impact on that day plus its impact on subsequent days. That is, it is the sum of $\beta_0 + \ldots + \beta_q$ (Schwartz 1993). To see this more easily, note that equation (1) can be recast as:

$Y_t = \alpha + \beta^*(w_0 X_t + \ldots + w_q X_{t-q}) + \varepsilon_t$

where the w_i are weights that sum to one, and β^* is $\beta_0 + \ldots + \beta_q$. That is, β^* is also interpretable as the marginal effect of a unit increase in a weighted average pollution variable. Since a unit increase in pollution on a single day increases the weighted average on all q subsequent days, the effect of

that single day's increase will be $\beta^* w_i$ on each of the q subsequent days, or β^* overall.

REANALYSIS APPROACH

In this reanalysis we have chosen the degree of freedom to be the same as in our previously published models where we used LOESS in GAM (Schwartz 2000a,b; Braga et al 2001) in order to enhance comparability. The autocorrelation of the residuals of the new models was still white noise, and in some cities autoregressive terms were added as in our original models. Two sensitivity analyses were performed.

To address the question of sensitivity of the effect size estimates to the choice of convergence criteria and maximum iterations in GAM, we refit the originally reported model changing only the convergence criteria. The stricter convergence criteria were the ones recommended by Dominici and coworkers (Dominici et al 2002), that is 10^{-15}.

In my second set of reanalyses, I used an approach where correct standard errors have been estimated. This has recently been done for the GAM models (Dominici, personal communication) but was not available in time for these analyses. The alternative approach involved either natural splines or penalized spline models.

The idea of a spline regression is simple. Suppose we want to smooth daily deaths or hospital admissions against temperature, and other covariates to obtain flexible estimates of their effects on outcome. We then have a model such as:

$$\text{Cases}_t \sim \text{Poisson}(\exp\{\beta_0 + f(\text{temp}_t) \qquad (1)$$
$$+ g(\text{season}_t) + \ldots + \eta PM_{10}\})$$

We can estimate the smooth function f (and similarly with g, etc) as follows:

$$f(\text{temp}_t) = \beta_{f1}\text{temp}_t + \beta_{f2}\text{temp}_t^2 + \beta_{f3}\text{temp}_t^3$$
$$+ \sum_k u_k^f(\text{temp}_t - \kappa_k)_+^3 \qquad (2)$$

where K is the number of boundary points (knots) between intervals of temperature, κ_k are the locations of those points, and $(\text{temp}_t - \kappa_k)_+$ is defined as:

$$\begin{cases} \text{temp}_t - \kappa_k, \text{temp}_t > \kappa_k \\ 0, otherwise \end{cases}$$

The u_k^f are the coefficients of the terms that start each knot. So far, this is simply a piecewise cubic fit, or a cubic regression spline. Splines are polynomial functions that are

pieced together to model the association between variables (Wypij 1996). Natural splines are cubic regression splines (as above) constrained to be linear at the extreme intervals (that is, the 2nd and 3rd derivatives are zero at the extremes).

Regression splines can be sensitive to the number and location of the knot points. Moreover, adding one degree of freedom to the model changes the location of all the other knot points when using natural splines, which makes comparisons of model fit with different number of degrees of freedom difficult. One alternative is a penalized spline.

If we instead fit equation (2) using a penalized likelihood (that is, imposing a penalty proportional for example to the sum of the squares of the u_k) we constrain the changes in the u_k at each knot point, effectively reducing the degree of freedom of the piecewise fit.

The larger the parameter (called the *smoothing parameter*) that multiplies the penalty term, the smaller the changes in slope that will minimize this penalized likelihood, and hence the smoother the curve. The degree of smoothness (and hence the value of the parameter) can be prespecified, or estimated by using generalized cross validation.

Software to use penalized splines to fit multiple covariates to a health outcome, including logistic and Poisson regressions has been developed by Wood (2000), and is available in R (The Comprehensive R Archive Network: *http://cran.r-project.org/*). This software uses cubic splines or thin plate splines.

When a penalized regression spline is used, we choose more knots than the expected degree of freedom required, and we use a penalized likelihood to constrain the u_k and therefore reduce the degree of freedom. In the regression spline approach, we choose a number of knots based on the desired degrees of freedom and fit the u_k without constraint. When we increase the degree of freedom in a penalized spline model, the location of the knot points does not change, only the extent of the penalty. This enhances comparisons.

I have used natural splines or penalized splines for my second set of reanalyses. Again, the same degrees of freedom were used to control for season, temperature on the day of and day before admission, relative humidity (or dew point depending on the original paper) and barometric pressure. In addition, we controlled for day of the week (indicator variables).

To reanalyze Schwartz (2000a), I used PM_{10} as a linear term for the mean of lag 0 and 1. For the reanalysis of Schwartz (2000b) and Braga (2001), I analyzed quadratic and unconstrained distributed lag models to examine how

PM_{10} effects were distributed over different lags. In those models we included terms for PM_{10} on the same day and each day between one and five days prior to the admission to capture the potential for delayed effect of pollution. The unconstrained distributed lag model of order q is:

$$\text{Cases}_t \sim \text{Poisson}(\exp\{\beta_0 + f(\text{temp}_t)$$
$$+\, g(\text{season}_t) + \ldots + \eta_0 PM_{10}t + \eta_1 PM_{10}t - 1 \quad (3)$$
$$+ \ldots + \eta_q PM_{10}t - q\})$$

The overall impact of a unit change in exposure on one day on admissions over the next q days is given by the sum $\eta_0 + \eta_1 + \ldots + \eta_q$. Even if PM_{10} at time t is correlated with $PM_{10}t - 1 \ldots PM_{10}t - q$, both the η_q and their sum will be unbiased estimators of the effects at each lag and of the overall effects, although the correlation will make the estimates unstable in a single city. Because they are unbiased, combining results across cities will produce more stable unbiased estimates.

An alternative is to provide more stability by constraining the variability of the η_i. The most common approach is to constrain the shape of the variation of the η_q's with lag number to fit some polynomial function. That is, the polynomial distributed lag model (PDL(q,d) with q lags and degree d), is the model (3) above, subject to the restriction:

$$\eta_j = \sum_0^d \lambda_k j^k \quad (4)$$

This approach originated with Almon (1965). More detailed description of distributed lag models has been published previously (Zanobetti et al 2000a,b; Schwartz 2000b).

In the second stage I combined the city-specific coefficients $\hat{\eta}_i$, using the covariance matrix of the estimates and again, inverse variance or covariance weighted averaging.

RESULTS

Table 1 shows the characteristics of the cities that were studied. In Table 2, I show the results of reanalyzing many of the analyses presented in Schwartz (2000a). The overall results, season-specific results, and copollutant-adjusted results are shown for both the new convergence criteria and using natural spline models. Overall, there was little change with the stricter convergence criteria and slightly lower effect size estimates using the natural spline models. The effect of PM_{10} remained, and remained independent of gaseous copollutants.

Table 1. Characteristics of 10 US Cities

City	1990 Population	Deaths	PM_{10} (µg/m^3)	Dew Point	Temp (°F)	Pressure (mm Hg)
New Haven	804,219	20.4	28.6	40.1	50.5	29.8
Birmingham	651,525	19.1	34.8	51.7	62.4	29.4
Pittsburgh	1,336,449	63.3	36.4	41.2	52.1	28.8
Detroit	2,111,687	59.7	36.9	40.7	50.9	29.3
Canton	367,585	9.9	29.31	41	50.4	28.7
Chicago	5,105,067	133.4	36.5	39.8	50.3	29.3
Minneapolis/St Paul	1,518,196	32.3	27.5	35.5	46.3	29.1
Colorado Springs	397,014	6	27.1	28.9	48.9	24.0
Spokane	361,364	8.7	40.6	34.2	47.9	27.5
Seattle	1,507,319	29.3	32.5	43.9	52.5	29.6

Table 2. Association Between PM_{10} (mean of lag 0 and lag1) and Daily Deaths in 10 US Cities: Analyses Stratified by Season (warm vs cold)

Model	Old GAM		New Convergence		Natural Splines	
	Percentage Increase in Deaths	95% CI	Percentage Increase in Deaths	95% CI	Percentage Increase in Deaths	95% CI
Overall	0.67	0.52, 0.81	0.66	0.52, 0.80	0.55	0.39, 0.70
Summer Only	0.67	0.48, 0.86	0.68	0.49, 0.87	0.52	0.31, 0.73
Winter Only	0.66	0.45, 0.87	0.61	0.42, 0.83	0.58	0.35, 0.80
In Hospital	0.49	0.31, 0.68	ND		ND	
Out of Hospital	0.89	0.67, 1.10	ND		ND	
Days < 50 µg/m^3	0.87	0.62, 1.12	ND		ND	
Confounding by						
$\quad$ SO$_2$	0.57	0.25, 0.90	0.64	0.33, 0.95	0.55	0.20, 0.90
$\quad$ CO	0.90	0.42, 0.97	0.89	0.54, 1.24	0.72	0.32, 1.13
$\quad$ O$_3$	0.69	0.53, 1.26	0.69	0.43, 0.94	0.59	0.32, 0.88

ND = not done.

Table 3 shows the results of the reanalysis of the distributed lag models for all cause mortality (Schwartz, 2000b). The effect size estimates were slightly reduced using the stricter convergence criteria, and slightly more using the penalized spline model. They remain considerably higher than estimates using a single exposure day.

Table 4 shows the results of the reanalysis of the Braga et al paper. There was little change in the effect estimates using the new convergence criteria for either chronic obstructive pulmonary disease (COPD), myocardial infarction (MI), or cardiovascular deaths. There was a substantial drop in the estimated effect of the two-day mean PM_{10} on pneumonia deaths with either the new convergence criteria or the penalized spline model. This drop indicates that the original results were far from convergence. In contrast, for the unconstrained distributed lag model, there was little change in the pneumonia results, with one of the new analyses showing slightly larger, and one slightly smaller effects. So the convergence problem seemed only to affect the two-day mean results. In all cases, PM_{10} remained a significant predictor of pneumonia deaths.

Table 3. Distributed Lag Between Air Pollution and Daily Deaths in 10 US Cities

City	Old GAM Quadratic Distributed Lag	Old GAM No Constraint	New Convergence Quadratic Distributed Lag	New Convergence No Constraint	Penalized Splines Quadratic Distributed Lag	Penalized Splines No Constraint
New Haven	1.85 (0.72)	1.80 (0.79)	1.79 (0.58)	1.77 (0.79)	1.88 (0.64)	1.78 (0.85)
Birmingham	0.36 (0.50)	0.34 (0.53)	0.36 (1.16)	0.34 (0.53)	−0.65 (0.52)	−0.65 (0.72)
Pittsburgh	0.89 (1.04)	1.00 (0.31)	0.65 (0.22)	0.62 (0.29)	0.58 (0.24)	0.56 (0.33)
Detroit	1.53 (0.32)	1.75 (0.30)	1.65 (0.22)	1.33 (0.30)	1.15 (0.25)	1.14 (0.35)
Canton	1.61 (1.25)	1.72 (1.36)	1.58 (1.02)	1.70 (1.35)	1.68 (1.12)	1.79 (1.51)
Chicago	0.98 (0.26)	0.91 (0.27)	1.08 (0.22)	1.06 (0.30)	1.01 (0.19)	0.98 (0.26)
Minneapolis	2.08 (0.49)	2.01 (0.53)	2.03 (0.39)	1.97 (0.53)	2.10 (0.40)	2.05 (0.55)
Colorado Springs	1.94 (1.18)	1.75 (1.26)	1.85 (1.04)	1.66 (1.25)	1.79 (1.11)	1.68 (1.40)
Spokane	2.04 (0.34)	0.74 (0.43)	1.10 (0.55)	1.10 (0.59)	0.96 (0.57)	0.94 (0.72)
Seattle	1.46 (0.31)	1.46 (0.34)	1.45 (0.28)	1.45 (0.34)	1.37 (0.30)	1.37 (0.36)
Overall	1.41 (0.13)	1.29 (0.13)	1.23 (0.14)	1.13 (0.14)	1.04 (0.11)	1.03 (0.14)

Table 4. Distributed Lag Between PM_{10} and Cause-Specific Deaths in 10 US Cities: Percentage Increases in Daily Cause-Specific Deaths and 95% CI Due to a 10 μg/m^3 Increase in PM_{10} Levels [a]

	COPD %	COPD 95% CI		Pneumonia %	Pneumonia 95% CI		CVD %	CVD 95% CI		MI %	MI 95% CI	
Original Results												
Two-day mean	1.5	0.7	2.3	6.6	−4.1	17.3	0.8	0.6	1.0	0.7	0.3	1.1
Unconstrained distributed lag	1.7	0.1	3.3	2.7	1.5	3.9	1.0	0.6	1.4	0.6	0.0	1.2
Stricter Convergence Criterion												
Two-day mean	1.5	0.8	2.2	2.1	1.4	2.9	0.8	0.5	1.1	0.6	0.2	1.1
Unconstrained distributed lag	1.9	0.1	3.7	3.1	1.6	4.6	1.0	0.6	1.5	0.7	−0.2	1.5
Penalized Spline Models												
Two-day mean	1.1	0.3	1.9	1.4	0.5	2.3	0.8	0.5	1.0	0.7	0.2	1.1
Unconstrained distributed lag	1.4	−0.5	3.4	2.2	0.6	3.8	0.9	0.4	1.4	0.5	−0.4	1.5

[a] COPD = chronic obstructive pulmonary disease; CVD = cardiovascular disease; MI = myocardial infarction.

DISCUSSION

In this paper I present the results of reanalyses of the association between PM_{10} and daily deaths in 10 US cities, using alternative methods to those originally published. In general, the results are similar. PM_{10} is still associated with daily deaths, the association is spread over several days, distributed lags yield larger effect size estimates, the associations are not confounded by gaseous air pollutants, and the lags between exposure and death vary somewhat by cause of death. There are also only minor differences between GAMs with the more stringent convergence criteria and the results of regression or penalized regression splines. In general, the latter yield slightly smaller effect estimates.

These results are similar to most of the other reanalyses of time-series studies that have been reported. That is, the main conclusions about association, confounding, and effect modification have not been changed. In most cases there are modest changes in size of the effect estimates although in a few cases they increased, and in a few instances (notably the 90 cities NMMAPS mortality estimates) the change in effect size estimates were more substantial (Dominici et al 2002). Moreover, these results are similar to the results of the APHEA1 study (Katsouyanni et al 1997), a large multicity study that did not use GAM models to control for season and weather, and hence was not effected by the problems with the GAM software. This general stability provides confidence that the overall conclusions previously reached about this extensive literature should remain unchanged.

REFERENCES*

Almon S. The distributed lag between capital appropriations and expenditures. Econometrica, 33:178–196, 1965.

Braga ALF, Zanobetti A, Schwartz J. The lag structure between particulate air pollution and respiratory and cardiovascular deaths in ten US cities. J Occup Environ Med, 2001;43:927–33.

Cleveland WS and Devlin SJ. 1988. Robust Locally-weighted regression and smoothing scatter plots. J Am Stat Assoc 74:829–36.

Coull BA, Schwartz J, Wand MP. 2001. Respiratory health and air pollution: Additive Mixed Model Analysis. Biostatistics 2: 337–49.

*Bold type identifies publications containing the original analyses revised in this short communication report.

Daniels MJ, Dominici F, Samet JM, Zeger SL. 2000. Estimating particulate matter mortality dose response curves and threshold levels: an analysis of daily time series for the 20 largest US cities. Am J Epidemiol 152:397–406.

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On generalized additive models in time series studies of air pollution and health. Am J Epidemiol 156: 193–203.

Eilers PHC, Marx BD. 1996. Flexible smoothing with B-splines and penalties (with discussion). Statistical Science 11:89–121.

Hastie TJ and Tibshirani RJ. 1990. Generalized additive models. Chapman and Hall. New York, NY.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, LeTertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J,Schindler C, Schwartz J. 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiol 125: 521–31.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak B, Sunyer J, Bacharova L, Schouten JP, Ponka A, Anderson HR. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: Results from time series data from the APHEA project. BMJ 314: 1658–1663.

Ramsay T, Burnett R, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwartz J, Zanobetti A. 2000. The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Schwartz J. Air pollution and daily mortality in Birmingham, Alabama. Am J Epidemiol:137:1136–1147 (1993).

Schwartz J. 1994. Nonparametric smoothing in the analysis of air pollution and respiratory illness. Canadian J Stat 22: 471–487.

Schwartz J. 2000a. Assessing confounding, effect modification, and thresholds in the association between ambient particles and daily deaths. Environ Health Perspect 108:563–568.

Schwartz J. 2000b. The distributed lag between air pollution and daily deaths. Epidemiology 11: 320–326.

Schwartz J, Coull B. Control for confounding in the presence of measurement error in heirarchical models. Biostatistics. In press.

Thurston SW, Eisen EA, Schwartz J. 2002. Smoothing in survival models: an application to workers exposed to metalworking fluids. Epidemiology 13: 685–692.

Wood, SN. 2000. Modeling and smoothing parameter estimation with multiple quadratic penalties. JRSS B 62: 413–428.

Wypij D. 1996. Spline and smoothing approaches to fitting flexible models for the analysis of pulmonary function data. American Journal of Respiratory & Critical Care Medicine 154: S223–S228.

Zanobetti A, Schwartz J, Dockery DW. Airborne particles are a risk factor for hospital admissions for heart and lung disease. Environ Health Perspect, 2000; 108:1071–1077.

Zanobetti A, Wand MP, Schwartz J, and Ryan L. 2000b. Generalized Additive Distributed Lag Models: Quantifying Mortality Displacement. Biostatistics 1: 279–292.

ABBREVIATIONS AND OTHER TERMS

CI	confidence interval
GAM	generalized additive model
LOESS	locally weighted smoothers
NO_2	nitrogen dioxide
O_3	ozone
PM_{10}	particulate matter < 10 µm in median aerodynamic diameter
SO_2	sulfur dioxide

Daily Deaths Associated with Air Pollution in Six US Cities and Short-Term Mortality Displacement in Boston

Joel Schwartz

ABSTRACT

Particulate air pollution has been associated with daily deaths in hundreds of cities in the world in the last 10 years. These studies have used a variety of methods to control for season and weather, but many recent studies have used nonparametric smoothing because of its greater flexibility. Recently, it has been discovered that the software used in those studies had relatively lax convergence criteria. As a result, some of the studies may have used regression models that did not fully converge. This may result in inaccurate estimates of the parametric terms, specifically the coefficient of air pollution. In addition, the software uses an approximation to estimate standard errors, which may result in inaccurate estimates of the standard errors in individual studies.

Recently, a new procedure has been developed for estimating the standard errors accurately (Dominici and Hastie 2002) but has only just become available. As a result, some have questioned the results of the portion of air pollution studies that used generalized additive modeling (GAM*). A number of such studies used data from six US cities that were part of the Harvard Six Cities Study (Ferris et al 1979). The investigators used the monitoring data that was collected as part of that study, matched with death certificate data from the National Center for Health Statistics, to show that:

1. $PM_{2.5}$ (particles less than 2.5 µm in median aerodynamic diameter) was more strongly associated with daily deaths than coarse particles (particles between 2.5 and 10 µm in median aerodynamic diameter [$PM_{2.5-10}$]) (Schwartz et al 1996);

2. There were independent associations of $PM_{2.5}$ particles from coal combustion and traffic, but not from crustal sources and daily deaths (Laden et al 2000); and

3. Rather than short-term mortality displacement or harvesting, examination of the association between particles and deaths on increasing time scales indicated cumulative effects (Schwartz 2000).

This paper reports the results of sensitivity analyses repeating the original analyses with either stricter convergence criteria or alternative methods that have neither the convergence nor the standard error issue in controlling for season and weather.

INTRODUCTION

Airborne particulate matter is a mixture of liquid and solid material of varying size and chemical characteristics. Ambient air quality standards for particulate matter in the United States used to be based on measurements of inhalable particles with median aerodynamic diameters of 10 µm or less (PM_{10}). This size range was chosen to limit regulation to those particles small enough to penetrate beyond the upper airways. Within that size range, there are several broad classes of particles. Mechanically generated particles from agriculture, mining, road traffic, and related sources are generally larger than 2.5 µm in aerodynamic diameter. These are usually referred to as *coarse mass* (CM) particles. In contrast, particles resulting from combustion processes are generally less than 2.5 µm ($PM_{2.5}$) in aerodynamic diameter. New US Environmental Protection Agency (EPA) regulations limit these particles as well. This size range is generally referred to as *fine particles*. These fine particles readily infiltrate into residential buildings with indoor levels similar to levels immediately outside the structure. Thus population exposure to fine particle mass has a higher correlation with day to day ambient particle measures than coarse particle mass or reactive gaseous pollutant concentrations (Sarnat et al 2001).

---

* A list of abbreviations and other terms appears at the end of the section.

Some of the fine particles are produced directly by combustion processes. These are generally all less than 1 µm in aerodynamic diameter, with a substantial fraction less than 0.3 µm in aerodynamic diameter and are often called *primary combustion particles.* Secondary combustion particles are produced by the chemical reaction in the atmosphere of gaseous emissions from combustion processes. In the eastern United States, these are primarily sulfate particles, and in the western United States, primarily nitrates. These secondary particles are usually between 0.3 and 1 µm in aerodynamic diameter and thus differ both in size distribution and chemical composition from the primary combustion particles. Because these different classes of particles have different sources, different regulatory policies are required to reduce their concentrations. If the reported associations between particulate air pollution and daily mortality are due principally to one class, this has obvious public policy implications.

The distribution and deposition of particles in the lung varies substantially with particle size. Coarse particles have a higher probability of being deposited in the bronchial region. Fine particles have a higher probability of being deposited in the periphery of the lung, especially in the respiratory bronchioles and alveoli where their clearance is slow relative to particles deposited on airways.

MATERIALS AND METHODS

SIX CITIES STUDY

Air Pollution Data

As part of the Harvard Six Cities Studies (Ferris et al 1979), dichotomous virtual impactor samples were placed at 6 central residential monitoring sites: Watertown, Massachusetts; Kingston-Harriman, Tennessee; St Louis, Missouri; Steubenville, Ohio; Portage, Wisconsin; and Topeka, Kansas. Separate filter samples were collected of fine particles ($PM_{2.5}$) and of the coarse mass ($PM_{2.5-10}$) fraction. The 2.5 µm cutpoint for dichotomous samplers represents the aerodynamic diameter at which 50% of the particles are sent to the fine particle filter and 50% to the coarse mass filter. The cutoff is relatively sharp, but some larger particles, including some crustal material, will still be deposited on the fine fraction filter. In addition, there is crustal material between 1 and 2.5 µm, representing the low end of the coarse-mode particle size distribution.

Integrated 24-hour samples were collected at least every other day from 1979 until the late 1980s, with daily sampling during health survey periods. For fine and coarse particle samples, mass concentration was determined separately by β attenuation. With the exception of a period between October 1981 and January 1984 in all cities, elemental composition of fine and coarse mass was determined by x-ray fluorescence. Elemental composition was available on 97% of these samples. In the fine fraction, 15 elements were routinely found above the limit of detection: silicon (Si), sulfur (S), chlorine (Cl), potassium (K), calcium (Ca), vanadium (V), manganese (Mn), aluminum (Al), nickel (Ni), zinc (Zn), selenium (Se), bromine (Br), lead (Pb), copper (Cu), and iron (Fe).

Source Identification

In separate analyses for each city, we used specific rotation factor analysis to identify up to five common factors from the 15 specified elements. We specified a single element as the tracer for each factor and maximized the projection of these elements using the Procrustes rotation, a variant of the oblique rotation method.

We defined three sources of fine particles in all six cities: a silicon factor classified as soil and crustal material, a lead factor classified as motor vehicle exhaust, and a selenium factor representing coal combustion sources. In city-specific analyses, we also considered vanadium (fuel oil combustion), chlorine (salt), and selected metals (nickel, zinc, or manganese) as possible targets and sources. We identified five source factors for each city, except for Topeka, where we were only able to identify three positive predictors of total $PM_{2.5}$.

Daily Factor Scores

For each metropolitan area we calculated daily scores for each of the identified factors. Information on the relative concentration of each of the 15 elements in each factor is provided by the standardized scoring coefficients. Therefore, the score for each factor for each day was calculated by multiplying the normalized concentration of each element by the respective standardized scoring coefficient for that element and factor and then summing across these 15 products.

We recentered these daily factor scores by calculating a factor score for a hypothetical day on which all of the elemental concentrations were zero and then subtracting this minimum factor score from the daily factor score. To rescale the factor scores from the normalized scale to the mass scale (in $\mu g/m^3$), we regressed the total daily fine particle concentrations on the daily factor scores for all of the factors in separate regression models for each city and obtained the product of each factor score with its regression coefficient. Only sources that were significant

predictors of total fine particle mass ($P < 0.10$) were considered in the mortality analyses.

Meteorologic Data

We obtained meteorologic data from the National Center for Atmospheric Research including hourly measures of temperature, dew point, and precipitation from the National Oceanographic and Atmospheric Administration weather station nearest to each city. We calculated 24-hour mean values for temperature and dew point.

Mortality Data

We defined the six metropolitan areas in this study as the county containing the air pollution monitor and contiguous counties (Schwartz et al 1996). For the mortality analysis, each study area is identified by the name of its largest city (eg, Watertown as Boston; Kingston-Harriman as Knoxville; Portage as Madison). We extracted daily deaths for people who lived and died in the selected counties from annual detail mortality tapes (National Center for Health Statistics) for the time periods with fine particle measurements. After excluding all deaths due to accidents and other external causes (*International Classification of Diseases, Ninth Revision, Clinical Modification* [ICD9] codes 800–999), we analyzed the remaining total daily deaths in this paper.

ANALYTIC METHODS

ORIGINAL ANALYSES

We analyzed the data from Schwartz and colleagues (1996) using GAM. For seasonal control, we used a locally weighted smoother (LOESS) with a window of 100 days. For control for weather terms (temperature and dew point), we used a span of 50% of the data. Day of the week was controlled using indicator variables. Laden and associates (2000) used the same approach but increased the span for temperature and dew point to 0.8 to account for the lesser amount of data. Note above that elemental composition data was missing for two years. The analysis of Boston data in Schwartz (2000) is described below.

NEW ANALYSES

The new analyses of the Schwartz and colleagues paper began by repeating the GAM models described above but with new convergence criteria of 10^{-15} for both the GLM and backfitting step, and maximum iterations of 1000 for both steps.

The next set of reanalyses used a variety of spline methods to estimate the association between air pollution and daily deaths. *Regression splines* refers to a process for controlling for a covariate that is flexible, allowing for a nonlinear relation. Polynomials are an obvious choice, but because of their inherent symmetries, they may not be flexible enough. For example, using a polynomial to control for the effects of temperature implies that the same coefficients explain the increased risk with both high and low temperature. A more flexible alternative is to divide the range of temperature into two or more ranges and to fit separate polynomials in each range. One approach to that is shown below:

$$f(\text{temp}_t) = \beta_1 \text{temp}_t + \beta_2 \text{temp}_t^2 + \beta_3 \text{temp}_t^3$$
$$+ \sum_{k=1}^{K} u_k^f (\text{temp}_t - \kappa_k)_+^3$$

where K is the number of boundary points (knots) between intervals of x, κ_k are the locations of those points, called knots, and $(x_i - \kappa_k)_+$ is defined as:

$$\begin{cases} x_i - \kappa_k, x_i > \kappa_k \\ 0, otherwise \end{cases}$$

This is a cubic regression spline. B-splines are merely a transformation of this equation that has better numerical properties in the computer estimation of β. Natural splines, called *restricted cubic splines* by some authors, are a variant that has the restriction that the function be linear at the extremes of the data.

It has long been noted with piecewise constant fits that the shape of the dose response indicated by plotting the fit was sensitive to the cut points chosen for the different categories. This sensitivity has led to the widespread use of quartiles or quintiles to define the cut points, or knots. This removes the choice of cut point from the discretion of the investigator and prevents misrepresentation. It does not eliminate the sensitivity, however. Higher order piecewise polynomials share this sensitivity to the location (and number) of knot points. This can be reduced by increasing the number of categories, but using many categories results in curves that are too wiggly.

One approach advocated by Eilers and Marx (1996) to address this issue is called *penalized splines*. One chooses a fairly large number of categories (thereby reducing the sensitivity to knot location) and then constrains the coefficients within each interval by imposing a penalty for the wiggliness of the resulting curve. For example, for the temperature spline illustrated above, the penalty would be

proportional to the sum of the squares of the changes in slopes that occurred in the fitted model. Instead of minimizing the sum of squares error, or more generally the log likelihood, one minimizes the penalized likelihood, which is the sum of the usual log likelihood and a parameter times the sum of the squares of the slope changes. Specifically the constraint is

$$\sum_{k=1}^{K} u_k^2 < C$$

and the model can be shown to be equivalent to a ridge regression. Software to use penalized splines to fit multiple covariates to a health outcome, including in logistic and Poisson regressions, has been developed by Wood and is available in R (The Comprehensive R Archive Network: $http://cran.r\text{-}project.org/$). This software uses cubic splines or thin plate splines.

The sensitivity to spline spacing may be a particular problem in this data set where the sampling frequency for airborne particles was irregular (mostly every other day, with occasional gaps, and with occasional periods of daily sampling). This suggests that basic regression spline models should be treated with caution. Because of this, more flexible alternatives are preferable. This paper reports two such alternatives. First, penalized splines were used. These were fit by tuning the smoothing parameters to produce models with roughly the same degrees of freedom. Because with penalized spline models one specifies a smoothness penalty, and not directly a degree of freedom, this entails repeated iterations, which were stopped when the degrees of freedom for each term were within a few percentage points of the original GAM model for each term. The number of knots as set at 50 for time (to achieve 36 *df*) and 10 for each weather term (with about 3.6 *df* as the ultimate target). Because of the greater number of knots, and the use of constraints to shrink in poorly estimated knot coefficients, this approach has less sensitivity to the distribution of gaps. In addition, I used an alternative regression spline approach.

Thin plate splines are an alternative basis for regression splines to B-splines (which are the basis of natural splines). This alone gives some indication of the sensitivity of the results. In addition, the implementation of thin plate splines in R uses an algorithm to choose the optimal knot placement given the data (Wood 2003), rather than simple approaches such as quantiles. This optimization of knot placement should also reduce sensitivity to knot location in the resulting analyses. Regression splines only use integer degrees of freedom, and these were rounded to the nearest whole number from those in the

original model, resulting in 4 *df* each for temperature and dew point in each city, and either 36 or 37 *df* for season.

Finally, to determine the sensitivity of the less flexible cubic spline approach, we used both natural cubic splines and B-splines, with the number of knots chosen to match the degree of freedom of the original models. Since B-splines have two fewer constraints than natural cubic splines (they are not constrained to be linear at the extremes), they have two fewer knots. Therefore all of the knots are placed at different locations, which provides a further test of sensitivity.

A similar, but less exhaustive analysis of model sensitivity was used for the source apportionment paper (Laden et al 2000) with the reanalysis restricted to GAMs with new convergence criteria and penalized splines. The penalized splines preserved the same degree of freedom per year as in the models in Table 1 but fewer degrees of freedom overall, reflecting the fact that the Laden and associates paper analyzed roughly two fewer years of data because of the period with no elemental analysis.

Similarly, I only reran the coarse mass regression models of Schwartz and colleagues (1996) using penalized splines with the same degree of freedom as the original GAM models.

BOSTON MORTALITY DISPLACEMENT ANALYSIS

Some have argued that the deaths occur only a few days early in persons who are already dying. This is usually referred to as *harvesting*.

Alternatively, exposure over time intervals of weeks may have some additional cumulative effect that is not captured in the current studies. Finally, prospective cohort studies of particulate matter and daily deaths (Pope et al 1995) have reported substantially larger effects of long-term exposure than are indicated by the daily time series studies. Those authors suggest the difference may represent an effect of chronic exposure. This paper examines how the association between particulate air pollution and mortality and morbidity varies as the time scale of the exposure varies.

Short-Term Mortality Displacement or Harvesting

To understand harvesting, imagine the population is divided into three states: a large healthy pool that is not immediately at risk of death, a smaller pool at high risk of death, and death. Three transitions are possible into and out of the risk pool: illness (T1), recovery (T2), and death (T3). This is illustrated in Figure 1. Under steady-state conditions, there is equilibrium between the number of deaths each day (T3) and the daily net recruitment into the high-risk pool (T1
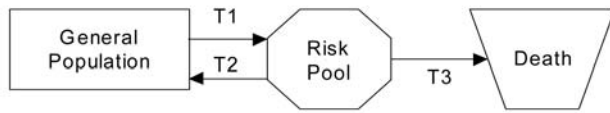
Figure 1. Schmatic of basic frality model with a single risk pool.

$- T2$). Any short-term increase in the mortality rate in the high-risk pool (R_3) that is not counterbalanced by an increase in the net recruitment rate (R_{1-2}) would deplete the high risk. This could occur if a transient environmental factor, such as a heat wave, increased mortality in the high risk pool but had little adverse effect on healthy subjects.

On subsequent days in response to that or another stressor, fewer people would die, because the mortality rate produced by the stressor would be applied to a smaller at-risk pool. This lower mortality rate would continue until the high-risk pool was replenished through net recruitment. As a result, the daily death count would be diminished on subsequent days. This is the *harvesting effect*.

If the stressor increases net recruitment into the risk pool by increasing the rate of serious respiratory or cardio-vascular illness, the harvesting effect would be diminished or possibly even reverse sign.

Analytic Approach

This paper used mortality data from Boston for the years 1979 through 1986, which has been described previously. Daily deaths were extracted from annual detail mortality tapes from the National Center for Health Statistics for Norfolk, Suffolk, and Middlesex Counties, which are the metropolitan counties proximate to the monitor. Deaths due to accidents and other external causes (ICD9 800–999) were excluded. Separate counts were also computed for deaths from ischemic heart disease (ICD9 410–414), pneu-monia (ICD9 480–486), and chronic obstructive pulmo-nary disease (COPD; ICD9 490–496). If air pollution only advances deaths by a few days, then we would expect an increase in daily deaths due to air pollution to be followed shortly by a decline. If we averaged over a week, the two effects would cancel out (or partially cancel out if some of the deaths are brought forward by a longer period).

A multiday average of daily deaths would no longer be associated with air pollution, because the air pollution effect and the rebound from it would have been smoothed over by the averaging. If we can separate the correlation between air pollution and daily deaths into characteristic averaging periods, the existence of an association at longer periods would demonstrate that all of the air pollution asso-ciated deaths are not being advanced by only a few days.

Cleveland's STL algorithm (seasonal and trend comp-osition using LOESS) (Cleveland 1990) was used to separate the time series of daily deaths into a long wavelength component (representing time trends and seasonal fluctuations, a midscale component, and a residual, very short timescale component. A window of 120 days was used to fit and remove the seasonal and long-term time trends in all analyses. Several different midscale components were examined (windows of 15, 30, 45, and 60 days). The $PM_{2.5}$ and weather data were similarly decomposed, and then log linear regressions were fit examining the association between $PM_{2.5}$ and daily deaths, independent of season and long-term time trends, for each of the different midscale frequency ranges. Temperature and dew point were controlled for in all regressions using nonparametric smoothing in a generalized additive model. The results for each of these windows were compared to results in the original regressions.

New Analysis

I repeated the analysis above using natural splines of temperature and dew point instead of nonparametric smooth functions in the log-linear models.

RESULTS

Table 1 shows the results of the new analyses of the Schwartz and colleagues 1996 paper. The first column shows the original results for comparison purposes. The second column shows the results of repeating the identical GAM models but with stricter convergence criteria. There is little change in the results. The remaining columns show the results using natural splines, B-splines, and thin plate splines, as well as the results using penalized splines.

In general, the results were quite comparable, with slightly lower estimates on average for the spline models compared to the GAMs (with new convergence criteria). The largest relative changes were seen in the two cities where the standard errors of the initial models were larger than the estimates.

Based on the broad consistency of these results, fewer approaches were used for reanalysis of the source apportionment study (Laden et al 2000). Table 2 shows the original results and the results when reanalyzed using GAM with stricter convergence criteria and when using penalized spline models with the same degree of freedom per year and per weather term. The overall pattern of results was unchanged in the reanalysis. The largest effect size estimate was for residual oil, followed by traffic and then coal. There was no association with dirt particles.

Table 1. Effect and Corresponding Standard Error of $PM_{2.5}$ (mean of Lag 0 and 1) on Daily Deaths in 6 US Cities

City	Original ($\beta \times 100$)	New Convergence	N-Splines	B-Splines	P-Splines	Thin Plate Splines
Boston	0.215 (0.034)	0.206 (0.034)	0.220 (0.038)	0.197 (0.038)	0.176 (0.039)	0.169 (0.039)
Kingston	0.138 (0.061)	0.121 (0.061)	0.120 (0.068)	0.112 (0.068)	0.102 (0.069)	0.087 (0.070)
Steubenville	0.101 (0.057)	0.094 (0.057)	0.068 (0.061)	0.059 (0.061)	0.071 (0.061)	0.055 (0.061)
St Louis	0.105 (0.033)	0.102 (0.033)	0.093 (0.035)	0.103 (0.035)	0.092 (0.035)	0.095 (0.035)
Madison	0.127 (0.077)	0.103 (0.077)	0.032 (0.085)	0.059 (0.085)	0.043 (0.086)	−0.005 (0.087)
Topeka	0.076 (0.142)	0.064 (0.143)	0.105 (0.158)	0.050 (0.156)	0.054 (0.160)	−0.010 (0.162)
Combined	0.146 (0.020)	0.137 (0.020)	0.129 (0.021)	0.117 (0.020)	0.113 (0.022)	0.104 (0.022)

Table 2. Effect ($\beta \times 100$) and Corresponding Standard Error of Source-Specific Particles on Daily Deaths in Six US Cities

City	Original	P-Splines
Boston		
Traffic	0.12 (0.16)	0.253 (0.174)
Coal	0.28 (0.08)	0.213 (0.094)
Residual oil	2.7 (1.3)	2.42 (1.58)
Knoxville		
Dirt	−0.17 (3.7)	−0.051 (1.06)
Traffic	0.52 (0.20)	0.381 (0.238)
Coal	0.08 (0.18)	0.087 (0.193)
St Louis		
Dirt	−0.30 (0.23)	−0.283 (0.259)
Traffic	0.43 (0.14)	0.416 (0.157)
Coal	0.030 (0.07)	0.027 (0.068)
Steubenville		
Dirt	−0.14 (0.29)	−0.115 (0.325)
Traffic	−0.20 (1.04)	−0.322 (1.17)
Coal	0.11 (0.12)	0.087 (1.29)
Residual oil	1.36 (2.56)	1.91 (2.67)
Madison		
Dirt	4.1 (3.4)	1.16 (3.70)
Traffic	0.63 (0.48)	1.00 (0.582)
Coal	0.09 (0.17)	−0.002 (0.192)
Topeka		
Dirt	−0.79 (1.8)	−0.349 (1.91)
Traffic	−0.82 (1.1)	−0.992 (1.18)
Coal	−0.39 (0.38)	−0.353 (0.428)
Combined		
Dirt	−0.23 (0.18)	−0.209 (0.198)
Traffic	0.34 (0.09)	0.355 (0.102)
Coal	0.11 (0.04)	0.079 (0.047)
Residual oil	2.43 (1.16)	2.29 (1.36)

Table 3. Effect and Corresponding Standard Error of Coarse Mass (Mean of > Lag 0 and 1) on Daily Deaths in Six US Cities

City	Effect	SE	*t*-statistic
Boston	0.000282	0.000533	0.53
Kingston	0.000669	0.000900	0.74
Steubenville	0.002043	0.001033	1.98
St Louis	0.000109	0.000496	0.22
Madison	0.000263	0.000972	0.27
Topeka	−0.00123	0.00109	−1.12

The magnitude of the associations were little changed with some increasing and some decreasing.

Table 3 shows the results of the penalized spline models for coarse mass in the six cities. As in the original analysis, coarse mass was only a significant predictor of daily deaths in Steubenville.

Table 4 shows the results of the reanalysis of the Boston Harvesting Study. The original results are presented for comparison. There was little change in the effect estimates at any scale and for any outcome, and again the same general pattern held. Effect estimates increased on progressively longer timescales except for COPD deaths, which showed evidence of harvesting.

Table 4. Effect and Corresponding Standard Error of Exposure to $PM_{2.5}$ on Daily Deaths on Different Time Scales

Scenario	Original Results	Natural Splines
All Deaths		
15 day	0.00226 (0.00026)	0.00227 (0.00027)
30 day	0.00256 (0.00027)	0.00253 (0.00027)
45 day	0.00304 (0.00028)	0.00304 (0.00029)
60 day	0.00368 (0.00027)	0.00369 (0.00028)
Ischemic Heart Disease Deaths		
15 day	0.00276 (0.00050)	0.00276 (0.00050)
30 day	0.00370 (0.00050)	0.00363 (0.00051)
45 day	0.00404 (0.00049)	0.00400 (0.00050)
60 day	0.00467 (0.00043)	0.00476 (0.00044)
Pneumonia Deaths		
15 day	0.00287 (0.00122)	0.00295 (0.00124)
30 day	0.00666 (0.00108)	0.00641 (0.00109)
45 day	0.01068 (0.00108)	0.01082 (0.00111)
60 day	0.01160 (0.00106)	0.01129 (0.00108)
Chronic Obstructive Pulmonary Disease Deaths		
15 day	0.00644 (0.00189)	0.00659 (0.00192)
30 day	0.00358 (0.00178)	0.00371 (0.00179)
45 day	0.00120 (0.00178)	0.00138 (0.00181)
60 day	−0.00050 (0.00166)	−0.00025 (0.00170)

DISCUSSION

The basic message from all of the original analyses remained unchanged with reanalysis using GAM with stricter convergence criteria, regression spline models, or penalized spline models. Often the size of the effect estimates was somewhat reduced using the spline models, but sometimes (eg, the effects of traffic particles on all deaths or of $PM_{2.5}$ on pneumonia deaths using a 15 day filter) they increased. The missing data problem combines with the sensitivity of seasonal fit to knot location in this data set to cause some sensitivity of the regression spline models to which splines and choices of knot placement are used. The penalized spline models have a clear advantage in this case and were used in the reanalysis of the Schwartz and colleagues (1996) and Laden and associates (2000) papers. The Boston harvesting paper used prefiltering to control for season in all cases and had only used smooth functions for weather terms. I refit those models using natural splines for convenience. These results add to the results of previous studies to suggest that the overall picture of air pollution time series studies is little changed by discovery of the problems in the GAM function. Regression splines do seem to sometimes jump around a bit in their results,

which is consistent with recent simulations by Dominici presented at the EPA GAM workshop (November 4–6, 2002, Research Triangle Park NC).

REFERENCES*

Cleveland RB, Cleveland WS, McRae JE et al. Seasonal-trend decomposition procedure based on LOESS. J Offic Stat 1990;6:3–73.

Dominici F, McDermott A, Zeger SL, and Samet JM. On the use of Generalized Additive Models in time-series studies of air pollution and health. Am J Epidemiol 2002; 156: 1–11.

Eilers PHC and Marx BD. Flexible smoothing with B-splines and penalties (with discussion). Statist Sci 1996; 89: 89–121.

Ferris BG Jr, Speizer FE, Spengler JD, Dockery D, Bishop YMM, Wolfson M, Humble C. Effects of sulfur oxides and respirable particles on human health: methodology and demography of populations in study. Am Rev Resp Dis 1979; 120: 767–779.

Hastie T, Tibshirani R. Generalized additive models. London, Chapman and Hall, 1990.

Laden F, Neas LM, Dockery DW, Schwartz J. Association of fine particulate matter from different sources with daily mortality in six US cities. Environ Health Perspect 2000;108:941–947.

Pope CA, Thun MJ, Naamboodiri MM et al. Particulate air pollution as a predictor of mortality in a prospective study of US Adults. Am J Resp Crit Care Med, 1995.

Sarnat JA, Schwartz J, Catalano PJ and Suh HH. Confounder or Surrogate: The Role of Gaseous Pollutants in Particulate Matter Epidemiology. Environ Health Perspect 2001;109:1053–1061.

Schwartz J, Dockery DW, Neas LM. 1996. Is daily mortality associated specifically with fine particles? J Air Waste Manage Assoc 46:927–939.

Schwartz J. Harvesting and long term exposure effects in the relation between air pollution and mortality. Am J Epidemiol 2000;151:440–448.

Wood S. Thin Plate Regression Splines. J Royal Statistical Society B, 2003, in press.

Wood SN. Modeling and smoothing parameter estimation with multiple quadratic penalties. JRSS B 2000; 62: 413–428.

* Bold type identifies publications containing the original analysis revised in this short communication report.

ABBREVIATIONS AND OTHER TERMS

CM	coarse mass
COPD	chronic obstructive pulmonary disease
EPA	US Environmental Protection Agency
GAM	generalized additive model
ICD9	*International Classification of Diseases, Ninth Revision, Clinical Modification*
LOESS	locally weighted smoother
PM_{10}	particles with median aerodynamic diameter of 10 µm or less
$PM_{2.5}$	particles less than 2.5 µm in median aerodynamic diameter
$PM_{2.5-10}$	coarse particles (between 2.5 and 10 µm in median aerodynamic diameter)

Ambient Air Pollution and Nonelderly Asthma Hospital Admissions in Seattle, Washington, 1987–1994

Lianne Sheppard

ABSTRACT

The establishment of air quality standards uses estimates from air pollution time-series studies. Recently evidence has surfaced that published effect estimates and confidence intervals (CIs*) from studies using generalized additive models (GAMs) may be biased due to premature convergence of the software and concurvity in the data. I reanalyzed data from a previously published report on nonelderly hospital admissions for asthma in Seattle, Washington, to evaluate the effect of the fitting procedure (Sheppard et al 1999). I found small changes in the results. Overall the effect estimates were slightly lower when more stringent convergence criteria were used in the GAM and an additional small reduction in the estimates was found when generalized linear models (GLMs) with natural splines were used instead. The CIs were slightly greater for the GLM fit. I also investigated the effect of not incorporating proper adjustment for imputing missing exposure data; estimates using singly imputed exposure data were slightly larger than the estimates that incorporated the uncertainty due to imputation. This bias was of the same order as the bias of using too liberal convergence criteria in the GAM. While all these biases were small, their importance was magnified because the effect estimates of interest were also quite small.

INTRODUCTION

Time-series regression models have been a mainstay of air pollution epidemiology because they have made it possible to estimate in large populations very small effects of ambient air pollutants on health. Recognized challenges of such models have included exposure measurement error

and difficulty in adequately controlling for the confounding effects of season and weather. Only recently has it become apparent that the computational tools used in these analyses, most notably GAM with nonprojection smoothers (eg, smoothing splines; see Hastie and Tibshirani 1990), have been producing poor estimates due to liberal default convergence criteria (Dominici et al 2002) and the effect of concurvity of the data (Ramsay et al 2003). The purpose of this paper is to report the reanalysis of a time-series study of nonelderly hospital admissions for asthma in the greater Seattle area (Sheppard et al 1999).

METHODS

ORIGINAL ANALYSIS

I briefly review the data and analysis of the original study, described by Sheppard and colleagues (1999). All asthma hospital admissions for people residing in 52 zip code areas in the greater Seattle area who were admitted to 23 hospitals were considered for this analysis. Asthma admissions were based on principal diagnosis; the small number of repeat admissions within 14 days were excluded. (In the original analysis, appendicitis admissions were also considered, as a control diagnosis.) Pollution and weather data were obtained from usual sources. When possible, daily measures were calculated by averaging over multiple sites. Large amounts of data for particulate matter less than 2.5 μm in diameter ($PM_{2.5}$) (72% and 81% per site) and for particulate matter less than 10 μm in diameter (PM_{10}) (4–40% per site) were lacking, as well as some sulfur dioxide (SO_2) measurements. With the exception of one year, ozone (O_3) was only measured seasonally. For PM and SO_2 but not O_3, we estimated missing data using multiple imputation. The high correlation between PM measurements and our access to a relatively complete series of light scattering measurements, also highly correlated with gravimetric PM, made it feasible to use gravimetric $PM_{2.5}$ and coarse mass in the analysis in spite of the large percentage of missing $PM_{2.5}$ data.

Analysis was conducted using S-Plus. Models were developed using a single dataset based on the average of the six imputed datasets. Models were selected for the degree of adjustment of seasonal and temperature confounding before the air pollution exposure data were included. Models were selected using three criteria: overall conceptual simplicity, small Akaike information criterion (AIC; Akaike 1970), and little evidence of underfitting or overfitting in the residual autocorrelation function. While these criteria are reasonable for selecting a good predictive model, there is no optimal procedure to assure appropriate control of confounding. The final model before introducing pollutants was a GAM with a log link, adjustment for day of week using indicator variables, and smoothing splines having 64 degrees of freedom (*df*) for time and 4 *df* for temperature.

Each pollutant was assessed in univariate analyses. Originally, model selection focused on the strongest effect for the one- to three-day lags of pollutants with additional exploration of same-day and longer lags as indicated. Because latency and induction times vary in a population, effect estimates that showed consistent patterns across lags were considered less likely to be spurious (ie, due only to model selection). Multipollutant models were based on pairs of pollutants from selected single-pollutant models. The linear pollutant effect was evaluated graphically, and season-specific and time period–specific models were considered. The final models showed small but consistent evidence of residual serial dependence in the final models, so the analyses were repeated using a GLM with natural splines to control for time and temperature in order to address the effect of residual dependence on the variance estimates using weighted adaptive variance estimates (Lumley and Heagerty 1999). These results were consistent with those obtained under the independence assumption.

REANALYSIS

In the reanalysis, I considered the asthma outcome only because the appendicitis outcome did not show consistent air pollution effects. (Appendicitis did significantly affect SO_2, but estimates from adjacent lags suggested this effect was probably spurious.) I fit the final models shown in Figures 3 and 4 of Sheppard et al (1999) for asthma using three analytic approaches: the original GAM with default convergence criteria (10^{-3} for both the backfitting and local scoring algorithms), the new GAM with stricter convergence criteria (10^{-8} for both the backfitting and local scoring algorithms), and a GLM with natural splines for time and temperature effects (with 64 and 4 *df* on equally spaced knots, respectively). Single pollutant analyses were performed on both the average dataset and the six separately

imputed datasets, allowing comparison of the results that incorporated adjustments for multiple imputation with the more common mean imputation approach.

RESULTS

The original paper showed the data distribution in Table 3 and correlation between predictors in Table 4. Briefly, asthma admissions averaged 2.7 per day with an interquartile range (IQR) of 3. Pollutant levels averaged 31.5 µg/m^3 for PM_{10}, 16.7 µg/m^3 for $PM_{2.5}$, 16.2 µg/m^3 for coarse PM, 30.4 ppb for O_3, 8 ppb for SO_2, and 1831 ppb for carbon monoxide (CO). The IQRs for these pollutants were 19, 11.8, 9.3, 20, 4.9, and 924, respectively. Large intercorrelations were found between the PM measurements and CO, ranging from 0.43 to 0.9. Temperature was negatively correlated with PM_{10}, $PM_{2.5}$, and CO, and positively correlated with O_3.

Figure 1 shows the results of the single-pollutant models. Solid lines show estimates after multiple imputation (if data were imputed) or after no imputation. For each section of the figure, the first of the three lines is the original estimate reported by Sheppard et al (1999). The second line is the estimate from the same model with stricter convergence criteria for backfitting and local scoring algorithms. The last line is the estimate from the

Figure 1. Percentage of change in hospital admission rates for asthma and 95% CIs for an interquartile range (IQR) increase in pollutants from single-pollutant models. Poisson regression models were adjusted for time trends (64-*df* spline), day of week, and temperature (4-*df* spline). Triplets of estimates for each pollutant represent the original GAM analysis using smoothing splines, the revised GAM analysis using stricter convergence criteria, and the GLM analysis using natural splines. For pollutants that required imputation, estimates ignoring (single imputation) or adjusting for (multiple imputation) the imputation are shown.

GLM with natural splines. For pollutants that required imputation, a second set of estimates, generated using single imputation based on the average of the six imputed datasets, are shown with dashed lines for the CIs.

Figure 1 shows a small but consistent decrease in estimates of the GAM model with stricter convergence criteria and an additional decrease when the GLM with natural splines was used. The width of the CIs of the estimates obtained with GLM also increased slightly. The bias of not adjusting for imputation of the data was of the same order of magnitude as the bias of using too liberal convergence criteria in the GAM. The estimated relative rates [95% CI] from the GAM with stricter convergence criteria were as follows: PM_{10} lagged 1 day, 1.04 [1.01–1.07]; $PM_{2.5}$ lagged 1 day, 1.04 [1.01–1.06]; coarse PM lagged 1 day, 1.02 [1.00–1.05]; CO lagged 3 days, 1.05 [1.02–1.08]; SO_2 same day, 1.01 [0.98–1.03]; and O_3 lagged 2 days, 1.07 [1.02–1.12]. The estimated relative rates [95% CI] from the GLM with natural spline smoothing were as follows: PM_{10} lagged 1 day, 1.03 [1.00–1.06]; $PM_{2.5}$ lagged 1 day, 1.03 [1.01–1.06], coarse PM lagged 1 day, 1.02 [0.99–1.04]; CO lagged 3 days, 1.04 [1.01–1.08]; SO_2 same day, 1.00 [0.97–1.04]; and O_3 lagged 2 days, 1.07 [1.01–1.13].

Figure 2 shows the results of two-pollutant models with both CO and $PM_{2.5}$, extending the results given in Figure 4 of the original paper. Estimates and 95% CIs are given for an IQR change in $PM_{2.5}$ lag 1 when CO is held constant (solid lines), for an IQR change in CO when $PM_{2.5}$ is held constant (dotted lines), and for half an IQR change in both pollutants simultaneously (dashed lines). For each of the three types of estimates, the three fitting procedures are shown from left to right as follows: the original GAM analysis, the GAM with stricter convergence criteria, and the GLM with natural spline smoothing. These results were generally consistent with Figure 1 and the original analysis. After rounding, the relative rate estimates for the joint change in $PM_{2.5}$ and CO (dashed lines) was 1.03 [95% CI = 1.01–1.05] for both the GAM with stricter convergence and the GLM with natural spline smoothing.

DISCUSSION

This reanalysis concentrated on replicating the results reported in Figures 3 and 4 of the original published analysis (Sheppard et al 1999). Overall the results did not change meaningfully. Small reductions in estimates occurred due to the alternate fitting procedures. The effect of single imputation (ie, not adjusting for replacing missing exposure data with an estimate of their expected values) increased the effect estimates slightly. In this dataset, this bias was of the same order as the bias of using too liberal convergence criteria in the GAM.

While the biases of computational details of the fitting were small, they were not completely trivial given the small effects of interest. Probably of more importance to this analysis and the literature in general is the bias, due to model selection, of reporting the most significant effects (Lumley and Sheppard 2003). Results of this analysis may be optimistic due to a model selection bias because the most statistically significant effects among three (and up to seven) models were reported for each pollutant. However, all the asthma effects showed a reasonable pattern of effects at adjacent lags, lessening but not removing the impact of model selection bias in this analysis. In simulations based on these data, Lumley and Sheppard (2000) showed that reporting the most significant of seven possible lags (without regard for the pattern of effects at adjacent lags) resulted in a median excess risk estimate of about half the estimated effect in the original study (1.5% for an 11.8 $\mu g/m^3$ change) when no underlying excess risk existed. The Monte Carlo P value for the simulation of the $PM_{2.5}$ effect in the original analysis was 0.08 (Lumley and Sheppard 2000), indicating that the original estimate was near the tail end but not outside the distribution of spurious excess risks from reporting the best of seven possible models.
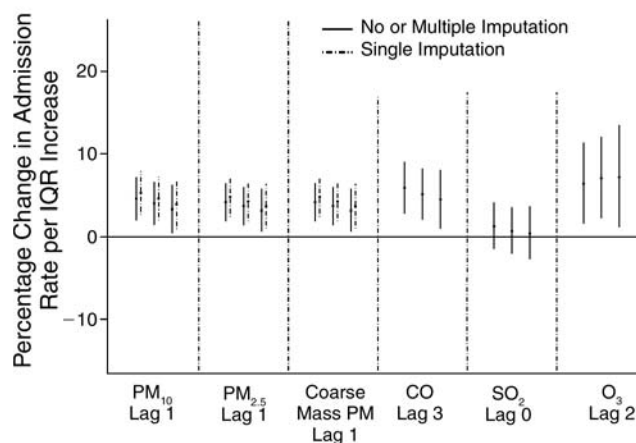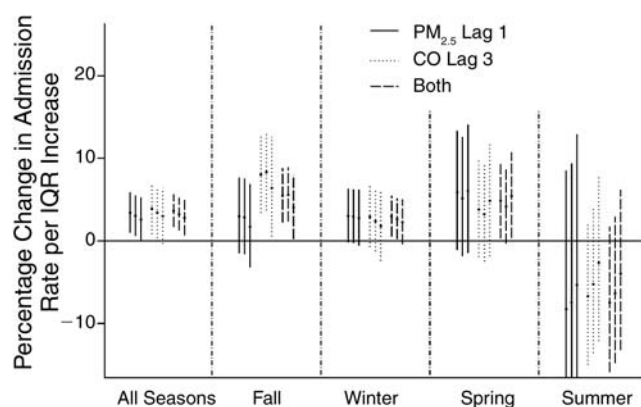

Figure 2. Percentage of change in hospital admission rates for asthma and 95% CIs for an interquartile range (IQR) increase in pollutants by season from a two-pollutant model (a change in both is half an IQR change in each). Poisson regression models were adjusted for time trends (64-*df* spline), day of week, and temperature (4-*df* spline). Triplets of estimates for each pollutant represent the original GAM analysis using smoothing splines, the revised GAM analysis using stricter convergence criteria, and the GLM analysis using natural splines.

REFERENCES*

Akaike H. Statistical predictor identification. Ann Inst Stat Math 1970; 22:203–217.

Dominici F, McDermott A, Zeger SL, Samet JM. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiology 2002; 156:193–203.

Hastie TJ, Tibshirani RJ. Generalized additive models. New York, NY: Chapman and Hall, Inc, 1990.

Lumley T, Heagerty PJ. Weighted empirical adaptive variance estimators for correlated data regression. J R Stat Soc Series B 1999;61:459–477.

Lumley T, Sheppard L. Assessing seasonal confounding and model selection bias in air pollution epidemiology using positive and negative control analyses. Environmetrics 2000; 11:705–717.

Lumley T, Sheppard L. Time series analyses of air pollution and health: Straining at gnats and swallowing camels? Epidemiology 2003; 14:13–14.

Ramsay TO, Burnett RT, Krewski D. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 2003; 14:18–23.

Sheppard L, Levy D, Norris G, Larson TV, Koenig JQ. Effects of ambient air pollution on nonelderly asthma hospital admissions in Seattle, Washington, 1987–1994. Epidemiology 1999; 10:23–30.

ABBREVIATIONS AND OTHER TERMS

CI	confidence interval
CO	carbon monoxide
df	degrees of freedom
GAM	generalized additive model
GLM	generalized linear model
IQR	interquartile range
O_3	ozone
PM_{10}	particulate matter less than 10 µm in diameter
$PM_{2.5}$	particulate matter less than 2.5 µm in diameter

---

* Bold type identifies publication containing the original analyses revised in this short communication report.

Daily Mortality and Fine and Ultrafine Particles in Erfurt, Germany

Matthias Stölzel, Annette Peters, and H-Erich Wichmann

ABSTRACT

This paper presents revised results from a study on short-term associations between air pollution and mortality from natural causes in Erfurt, Germany using more stringent convergence criteria than the default in the GAM* (generalized additive model) function of the statistical software package S-Plus.

The objective of the original study was to assess the association between particles of different size including ultrafine particles and mortality. Particle number concentrations (NCs) ranging in size from 10 nm to 2.5 µm were measured from September 1995 to December 1998 using a mobile aerosol spectrometer (MAS). Size-specific particle mass concentrations (MCs) were computed by assuming a mean density of particles. Death certificates were obtained from local health authorities. Infant deaths were excluded. The data were analyzed using Poisson regressions adjusting for trend, season, day of week, influenza epidemics, and meteorology.

For the reanalysis we used more stringent convergence criteria for the GAM function in S-Plus. Furthermore, we also fit a parametric generalized linear model (GLM) with the same degree of freedom as in GAM.

Different modeling approaches did not result in major changes of the effect estimates. The relative mortality risk for an increase of the ultrafine particle NC per interquartile range (IQR) was 1.045 (95% confidence interval [CI]: [0.996,1.096]) under the strict GAM and 1.047 (95% CI: [0.993,1.103]) under the GLM, respectively, compared to 1.046 (95% CI: 0.997,1.097]) under the original GAM. The respective numbers for fine particle mass were: 1.030 (95% CI: [0.999,1.062]) for the strict GAM, 1.029 (95% CI: [0.993,1.066]) for the GLM and 1.031 (95% CI: [1.000,1.063])

for the default GAM. The confidence intervals obtained under GLM are wider than under GAM indicating the underestimation of standard errors in the current GAM implementation.

The results of the original study do not seem to be sensitive to the selection of convergence parameters of the GAM function in S-Plus or to alternative modeling approaches.

INTRODUCTION

Increases in morbidity and mortality have been observed consistently and coherently in association with ambient air pollution. A number of studies on short-term effects have identified ambient particles as a major pollutant in urban air (Schwartz 1994, Katsouyanni et al 1997, Pope and Dockery 1999). Particles with smaller diameters such as fine particles ($PM_{2.5}$: particulate matter with a diameter < 2.5 µm) seem to have a stronger association with mortality than inhalable particles (PM_{10}: particulate matter with a diameter < 10 µm) (Schwartz et al 1996).

A number of studies on the associations between adverse health outcomes and elevated concentrations of particulate matter use the GAM (Hastie and Tibshirani 1990) function in the statistical software package S-Plus (Insightful Corporation, Seattle WA) for their regression analyses. Recently, there were some concerns about inadequate convergence criteria for this function (Dominici et al 2002) as well as about the underestimation of standard errors with this function. The purpose of this paper is to present reanalyses of the regression results of a study that was conducted in Erfurt, Germany (Wichmann et al 2000), using more stringent convergence criteria and an alternative modeling approach to address the standard error problem.

The objective of the original study (Wichmann et al 2000) was to assess the association and role of particles of different size (with size classes ranging from 10 nm to 2.5 µm) with respect to mortality in an urban setting. One specific aim was to compare the role of ultrafine and fine particles. Ultrafine particles have a diameter less than

---

100 nm. They mainly contribute to the NC of particles and contribute almost nothing to the MC. It has been hypothesized that ultrafine particles contribute to the associations between particulate matter and health outcomes (Oberdörster et al 1995, Seaton et al 1995]). Ultrafine particles have been associated with decreases in lung function, increased symptom prevalence, and increased medication use in adult asthmatics (Peters et al 1997, von Klot et al 2002). Recent epidemiologic findings on health effects of ultrafine particles have been reviewed by Wichmann and Peters (2000) and Ibald-Mulli et al (2002).

In the original study (Wichmann et al 2000), which is reanalyzed here, a 4.6% increase in daily deaths from natural causes was observed when the ultrafine particle NC increased over its IQR (95% CI: $-0.3, 9.7$). The respective increase for fine particle mass was 3.1% (95% CI: 0.0, 6.3). Whereas the association between particle mass and mortality was observed on the same day of exposure the ultrafine particle number association was delayed. The strongest association was observed on day 4 after exposure.

METHODS

The methods of this study have been described in detail by Wichmann et al (2000). Briefly, the study was conducted in Erfurt, a medium-sized town in former Eastern Germany with about 200,000 inhabitants. The central monitoring site was located 1 km south of the city center and 40 m west of a major road. An MAS was run to measure particle NCs in different size fractions continuously, combining an electrical mobility analyzer and a laser spectrometer covering the size range of particles with diameters of 10 nm to 2.5 μm. Size-specific MCs were computed by assuming a mean density of particles. In the final analyses, the daily mean values were used. In the following, NC_{a-b} and MC_{a-b} denote the NC and MC of particles with a diameter between a and b μm, respectively.

Furthermore, sulfur dioxide (SO_2) was measured at this site. PM_{10} and $PM_{2.5}$ were collected on a Harvard impactor. Data on nitrogen dioxide (NO_2) and carbon monoxide (CO) were obtained from a state-run station (Krämpferstraβe), which was only 2 km north of the GSF measurement station. Death certificates were obtained from local health authorities and aggregated to time series of counts for total mortality.

The statistical methods used in this study have been described in detail by Wichmann et al (2000). Briefly, the daily mortality counts were analyzed using Poisson regression models. Trend, season, day of week, influenza epidemics, and meteorology were considered potential confounders. Because the shape of the dose–response relation for most confounders is unknown, nonparametric smooth functions like smoothing splines or locally weighted regression smoothers (LOESS) were used to adjust for them. This is possible in GAM (Hastie and Tibshirani 1990). Model building was done with respect to visual inspection of the plotted fit and the Akaike information criterion (AIC). Furthermore, we tried to remove as much autocorrelation and overdispersion as possible. In detail:

1. Trend and season were modeled as a smoothing spline with 14 degrees of freedom (*df*) (ie, approximately 4 *df*/year). The degrees of freedom were chosen by visual inspection of the plot. Variations in the size, shape, and placement of a winter maximum each year were permitted, but several extra maxima per year were not permitted.
2. Day of week was included as an indicator variable.
3. Influenza data were obtained from an external source consisting of weekly reports of acute respiratory cases. Strong epidemics, which were visible to the naked eye in the raw mortality time series, turned out to be associated with mortality. They were included with a LOESS term as a confounder. The data were only available for the whole of Germany, but the epidemics may have hit Erfurt at a different time than the rest of Germany. Therefore, the best shift was decided by the AIC.
4. Meteorology: Immediate and delayed temperature effects were incorporated in the confounder model by LOESS terms with the span chosen so that the shape was either monotonous or concave or convex. Humidity effects were also included as a LOESS term.

After fitting the confounders the pollutant variables were fitted parametrically. Two transformations were considered for each pollutant: identical (id) transformation and log transformation. Several lags up to lag 5 were tested for each variable allowing each to select its best fitting day ("Best single-day lag").

In the initial study we used the GAM function in S-Plus (Insightful Corp) with the default convergence criteria (`epsilon = 10⁻³, bf.epsilon = 10⁻³, maxit = 10, bf.maxit = 10`). The first two parameters control the convergence of the local scoring algorithm and the backfitting algorithm, respectively. The latter two limit the number of iterations in the respective algorithm. For the reanalyses we used the following settings: `epsilon = 10⁻¹⁵, bf.epsilon = 10⁻¹⁵, maxit = 1000, bf.maxit = 1000`.

To address the problem of underestimation of standard errors in the GAM function of S-Plus, we also fitted a GLM (McCullagh and Nelder 1989) where all the confounders mentioned above were included as parametric smooth functions with the same degree of freedom as in the original GAM. Natural splines were chosen as smooth functions except for the influenza epidemics, which were modeled as polynomials.

RESULTS

The shape of the dose–response relation for the confounders looked very similar for the GAM with smoothing splines and LOESS terms and the GLM with natural splines and polynomials (Figure 1). In general, the natural spline is a bit smoother than the LOESS term. Furthermore, the natural spline smooths fit extreme maxima and minima better.

Table 1 shows regression results for the best single-day lags for NCs and MCs of particles in different size classes and for gaseous pollutants that were computed with three different approaches: GAM with default convergence criteria as published by Wichmann et al (2000) (*default*); GAM reanalyzed with more stringent convergence criteria (*strict*); and a fully parametric GLM with natural splines and polynomials (*GLM*). The results are given as relative risks (RRs) per IQR. Transformation of the pollution variable (log or id transformation) is indicated in the transformation (TR) column of the table.

In general, the RR estimates do not differ much throughout the three approaches. The key findings of the original study for the NC of ultrafine particles ($NC_{0.01-0.1}$) and the MC of fine particles ($MC_{0.01-2.5}$) remain almost unaffected (Table 1).

In the GLM with natural splines, the risk estimates tend to be somewhat smaller than their counterparts in the two GAMs for most of the pollutants except for NCs of particles. Furthermore, the confidence intervals here are wider than in the GAM and therefore affect the statistical significance of some of the risk estimates. This was in particular observed for fine particle MC ($MC_{0.01-2.5}$) and PM_{10}.

The same holds true when taking into account all lags and all transformations for a selected group of pollutants (Table 2). Again, there is almost no change in the risk estimate when using GAM with more stringent convergence criteria instead of the default criteria. Also, the estimates computed for GLM do not differ much with the exception of $MC_{0.01-2.5}$, PM_{10}, NO_2, and CO each at lag 0 with log transformation and NO_2 and CO each at lag 1 with id transformation where the RR estimate for the GLM is up to 0.019 smaller than the risk estimate obtained for the GAM with strict convergence criteria. Again, the confidence intervals in the fully parametric model (GLM) are in general wider than in the two GAMs.

When computing two-pollutant models without and with interaction (Tables 3 and 4, respectively), then again almost no differences between the two GAMs with the different convergence criteria could be found. There was a tendency toward marginally smaller effect estimates (maximal difference, 0.002) in the case of two pollutant models with interaction when using GAM with more stringent convergence criteria (Table 4).

In the GLM without interaction, the effect estimates for the ultrafine particle NC ($NC_{0.01-0.1}$) remain fairly constant in the presence of other copollutants. The relative risk estimates for $NC_{0.01-0.1}$ are in the range of 1.042 to 1.048 and therefore similar to the one-pollutant model where the RR = 1.047. When using GAM, those effect estimates were in general somewhat smaller (1.034–1.044 in comparison to 1.045 in the one-pollutant model).
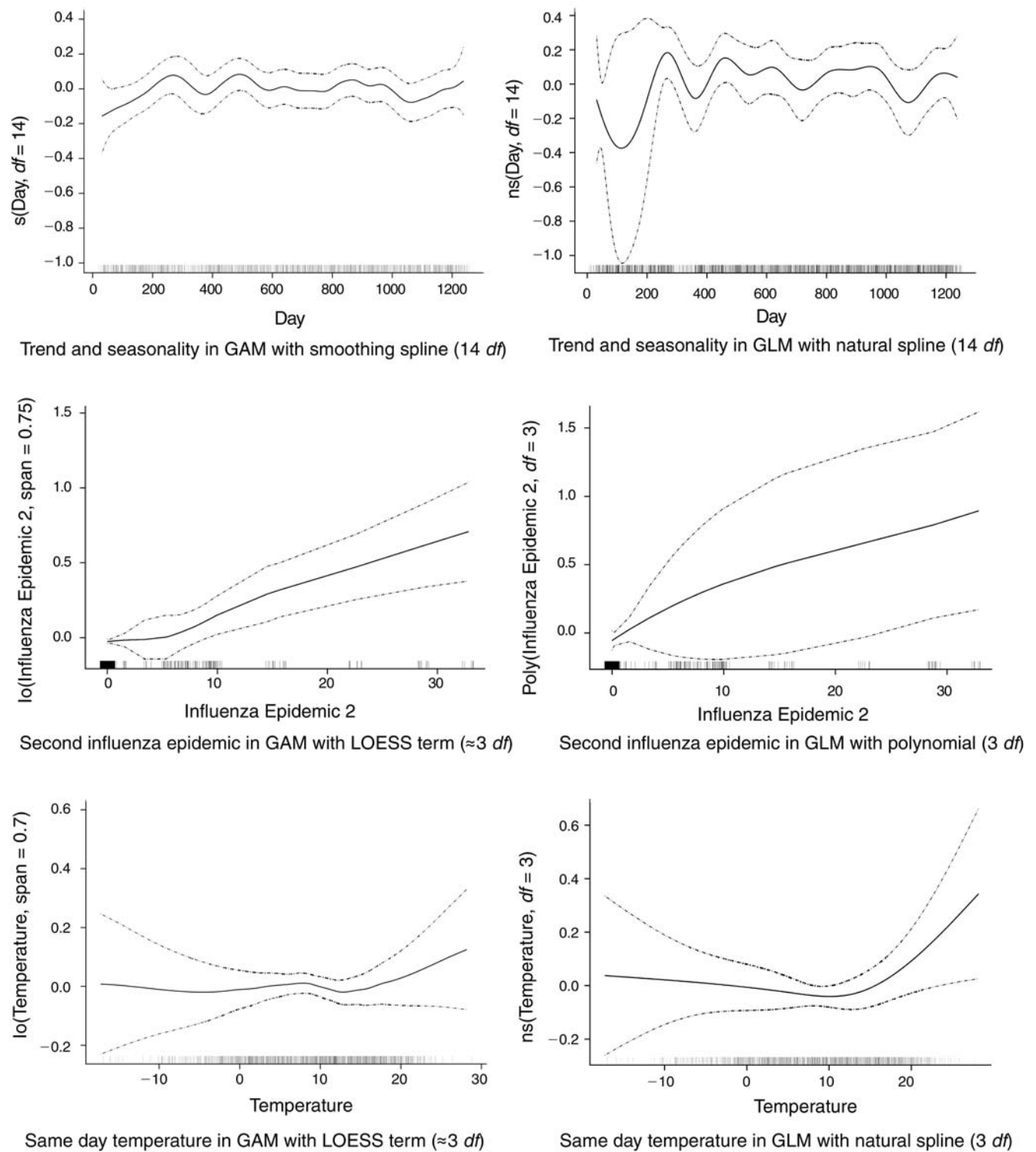
Figure 1. Dose–response relations for some confounders.

Table 1. Regression Results for Daily Mortality for Particle Number Concentration, Particle Mass Concentration, and Gaseous Pollutants. Reanalysis of Table 18 in Wichmann et al (2000, p 38)[a]

	Lag (Days)	TR	IQR	Default			Strict			GLM		
				RR	CI	P	RR	CI	P	RR	CI	P
Particle Number Concentration (Particles/cm³): Best Single-Day Lag												
$NC_{0.01-0.03}$	4	log	5177–14065	1.048	1.000 1.099	0.05	1.047	0.999 1.098	0.06	1.049	0.995 1.105	0.07
$NC_{0.03-0.05}$	4	id[b]	1603–4127	1.031	0.998 1.066	0.07	1.031	0.998 1.065	0.07	1.034	0.999 1.071	0.06
$NC_{0.05-0.10}$	1	log	993–2518	1.043	0.999 1.089	0.06	1.043	0.999 1.089	0.06	1.039	0.991 1.090	0.11
$NC_{0.01-0.10}$	4	log	8042–20732	1.046	0.997 1.097	0.07	1.045	0.996 1.096	0.07	1.047	0.993 1.103	0.09
$NC_{0.01-2.5}$	4	log	9659–22928	1.041	0.991 1.093	0.11	1.040	0.991 1.092	0.11	1.043	0.989 1.101	0.12
Particle Mass Concentration (µg/m³): Best Single-Day Lag												
$MC_{0.1-0.5}$	0	id	9.8–25.2	1.026	0.995 1.058	0.10	1.026	0.995 1.058	0.10	1.024	0.988 1.062	0.19
$MC_{0.5-1.0}$	0	id	0.81–4.03	1.015	0.996 1.034	0.13	1.015	0.996 1.034	0.13	1.013	0.992 1.034	0.24
$MC_{1.0-2.5}$	3	id[c]	0.56–1.55	0.977	0.954 1.001	0.06	0.977	0.954 1.001	0.06	0.978	0.953 1.004	0.10
$MC_{0.01-}$	0	id	11.3–31	1.028	0.996 1.060	0.09	1.028	0.996 1.060	0.09	1.024	0.988 1.062	0.20
$MC_{0.01-2.5}^{0.10}$	0	id	12–31.9	1.031	1.000 1.063	0.05	1.030	0.999 1.062	0.05	1.029	0.993 1.066	0.12
Other Particle Mass (µg/m³): Best Single Day Lag												
$PM_{2.5}$	3[d]	id	13–31.5	0.970	0.941 1.000	0.05	0.970	0.941 0.999	0.05	0.975	0.941 1.010	0.16
PM_{10}	0	id	19.9–47.6	1.035	1.001 1.069	0.04	1.034	1.001 1.069	0.04	1.029	0.990 1.069	0.14
TSP	1	log	28.8–61.9	1.023	0.981 1.067	0.28	1.023	0.981 1.067	0.29	1.014	0.968 1.062	0.56
Gaseous Pollutants: Best Single-Day Lag												
SO_2	0	log	5.5–19.8	1.060	1.011 1.112	0.02	1.061	1.011 1.112	0.02	1.058	1.003 1.116	0.04
NO_2	4	id	26–46	1.029	0.992 1.067	0.12	1.029	0.991 1.067	0.14	1.025	0.983 1.068	0.24
CO	4	log	0.3–0.8	1.055	1.003 1.110	0.04	1.054	1.003 1.109	0.04	1.047	0.984 1.114	0.15

[a] IQR = interquartile range; TSP = total suspended particles; id = identical.

[b] With log transformation fit was only slightly less well: lag = 4, TR = log, RR = 1.040, CI = [0.994–1.089], P = 0.09 for epsilon=10^{-3}, RR = 1.040, CI = [0.993–1.088], P = 0.09 for epsilon = 10^{-15}, and RR = 1.041, CI = [0.991-1.094], P = 0.11 for GLM, respectively.

[c] The second best fit was lag = 0, TR = id, RR = 1.019, CI = [0.997–1.042], P = 0.10 for epsilon=10^{-3}, RR = 1.019, CI = [0.997–1.042], P = 0.10 for epsilon = 10^{-15}, and RR = 1.016, CI = [0.992–1.041], P = 0.19 for GLM, respectively.

[d] The second best fit was lag = 0, TR = id, RR = 1.019, CI = [0.991–1.049], P = 0.19 for epsilon = 10^{-3}, RR = 1.019, CI = [0.991–1.049], P = 0.19 for epsilon = 10^{-15}, and RR = 1.015, CI = [0.981–1.049], P = 0.40 for GLM, respectively.

Table 2. Regression Results for Total Mass Concentration (MC), Number Concentration (NC), and Gaseous Pollutants: Total Mortality, All Single-Day Lags and Transformations. Reanalysis of Table 19 in Wichmann et al (2000, p 40)[a]

Lag (days)	TR	Default RR/IQR	Default CI	Default P	Strict RR/IQR	Strict CI	Strict P	GLM RR/IQR	GLM CI	GLM P
NC$_{0.01-0.10}$										
0	id	1.022	0.982 1.065	0.22	1.023	0.982 1.065	0.28	1.023	0.979 1.069	0.31
0	log	1.019	0.969 1.072	0.30	1.019	0.969 1.072	0.45	1.019	0.965 1.077	0.50
1	id	1.003	0.966 1.042	0.39	1.003	0.966 1.042	0.88	1.004	0.963 1.046	0.85
1	log	1.026	0.979 1.075	0.22	1.025	0.979 1.074	0.29	1.025	0.974 1.078	0.34
2	id	0.984	0.946 1.022	0.28	0.983	0.946 1.022	0.39	0.987	0.947 1.029	0.55
2	log	0.994	0.948 1.042	0.39	0.993	0.948 1.041	0.78	0.995	0.946 1.047	0.86
3	id	1.009	0.970 1.050	0.36	1.009	0.969 1.050	0.67	1.014	0.971 1.060	0.52
3	log	1.021	0.973 1.071	0.28	1.020	0.973 1.070	0.41	1.024	0.972 1.079	0.37
4	id	1.035	0.995 1.077	0.10	1.034	0.994 1.076	0.10	1.038	0.994 1.083	0.09
4	log	1.046	0.997 1.097	0.07	1.045	0.996 1.096	0.07	1.047	0.993 1.103	0.09
5	id	1.005	0.967 1.044	0.39	1.004	0.967 1.043	0.84	1.004	0.964 1.046	0.83
5	log	1.012	0.967 1.059	0.35	1.011	0.967 1.058	0.62	1.011	0.963 1.062	0.66
MC$_{0.01-2.5}$										
0	id	1.031	1.000 1.063	0.06	1.030	0.999 1.062	0.05	1.029	0.993 1.066	0.12
0	log	1.040	0.992 1.089	0.11	1.039	0.992 1.088	0.11	1.027	0.975 1.082	0.31
1	id	1.013	0.982 1.045	0.29	1.013	0.982 1.044	0.43	1.012	0.978 1.048	0.49
1	log	1.016	0.969 1.064	0.32	1.014	0.968 1.063	0.55	1.006	0.956 1.059	0.81
2	id	1.000	0.969 1.032	0.40	1.000	0.968 1.032	0.98	1.003	0.967 1.040	0.86
2	log	0.998	0.951 1.047	0.40	0.997	0.951 1.046	0.91	0.995	0.944 1.049	0.86
3	id	0.978	0.947 1.009	0.15	0.977	0.947 1.009	0.15	0.982	0.946 1.020	0.35
3	log	0.985	0.939 1.032	0.32	0.984	0.939 1.031	0.49	0.983	0.933 1.037	0.53
4	id	1.004	0.974 1.035	0.39	1.003	0.974 1.034	0.83	1.009	0.976 1.043	0.59
4	log	1.006	0.962 1.053	0.38	1.005	0.961 1.052	0.82	1.005	0.957 1.056	0.84
5	id	1.006	0.977 1.037	0.37	1.006	0.976 1.037	0.70	1.011	0.979 1.044	0.51
5	log	1.023	0.977 1.071	0.25	1.022	0.976 1.070	0.35	1.024	0.975 1.074	0.35
PM$_{10}$										
0	id	1.035	1.001 1.069	0.05	1.034	1.001 1.069	0.04	1.029	0.990 1.069	0.14
0	log	1.030	0.987 1.075	0.16	1.030	0.987 1.075	0.17	1.017	0.969 1.068	0.49
1	id	1.016	0.984 1.050	0.25	1.016	0.984 1.050	0.33	1.015	0.978 1.053	0.44
1	log	1.020	0.978 1.064	0.26	1.020	0.978 1.064	0.35	1.014	0.968 1.063	0.56
2	id	0.997	0.964 1.030	0.39	0.996	0.964 1.030	0.84	0.998	0.961 1.036	0.91
2	log	0.993	0.951 1.036	0.38	0.993	0.951 1.036	0.73	0.989	0.943 1.037	0.64
3	id	0.976	0.944 1.010	0.15	0.976	0.943 1.009	0.15	0.979	0.941 1.018	0.28
3	log	0.988	0.947 1.031	0.34	0.988	0.946 1.031	0.58	0.986	0.939 1.035	0.57
4	id	1.003	0.971 1.035	0.39	1.002	0.971 1.034	0.89	1.007	0.972 1.043	0.71
4	log	1.002	0.962 1.044	0.40	1.002	0.961 1.044	0.94	1.003	0.959 1.049	0.91
5	id	0.995	0.964 1.028	0.38	0.995	0.963 1.027	0.74	0.998	0.964 1.033	0.90
5	log	1.007	0.966 1.049	0.38	1.006	0.965 1.048	0.78	1.007	0.964 1.052	0.75

(Table continues next page)

[a] IQR = interquartile range; TR = transformation; id = identical.

Table 2 (Continued). Regression Results for Total Mass Concentration (MC), Number Concentration (NC), and Gaseous Pollutants: Total Mortality, All Single-Day Lags and Transformations. Reanalysis of Table 19 in Wichmann et al (2000, p 40)[a]

Lag (days)	TR	Default			Strict			GLM		
		RR/IQR	CI	P	RR/IQR	CI	P	RR/IQR	CI	P
SO₂										
0	id	1.020	0.994 1.046	0.13	1.020	0.994 1.047	0.14	1.023	0.991 1.056	0.17
0	log	1.060	1.011 1.112	0.02	1.061	1.011 1.112	0.02	1.058	1.003 1.116	0.04
1	id	0.993	0.967 1.019	0.35	0.993	0.967 1.019	0.60	0.996	0.966 1.028	0.81
1	log	1.019	0.974 1.067	0.29	1.019	0.973 1.067	0.42	1.018	0.967 1.071	0.50
2	id	1.020	0.994 1.047	0.13	1.020	0.994 1.047	0.14	1.025	0.994 1.057	0.12
2	log	1.019	0.973 1.068	0.29	1.019	0.972 1.068	0.43	1.020	0.969 1.074	0.45
3	id	1.008	0.982 1.034	0.33	1.008	0.982 1.034	0.55	1.015	0.983 1.047	0.36
3	log	1.049	1.002 1.098	0.05	1.048	1.001 1.098	0.04	1.051	0.999 1.106	0.06
4	id	1.000	0.975 1.025	0.40	0.999	0.974 1.025	0.96	1.005	0.976 1.035	0.73
4	log	1.000	0.956 1.046	0.40	0.999	0.955 1.045	0.97	1.002	0.955 1.052	0.93
5	id	1.016	0.991 1.042	0.18	1.016	0.991 1.042	0.21	1.020	0.993 1.049	0.15
5	log	1.038	0.992 1.086	0.11	1.038	0.992 1.086	0.11	1.041	0.992 1.092	0.10
NO₂										
0	id	1.003	0.965 1.043	0.39	1.003	0.964 1.044	0.89	0.991	0.949 1.036	0.70
0	log	0.989	0.951 1.029	0.35	0.988	0.949 1.029	0.57	0.975	0.932 1.020	0.27
1	id	1.015	0.977 1.054	0.30	1.015	0.976 1.054	0.46	1.006	0.964 1.049	0.79
1	log	1.020	0.982 1.059	0.24	1.019	0.981 1.059	0.33	1.008	0.966 1.052	0.71
2	id	1.004	0.966 1.044	0.39	1.003	0.964 1.044	0.87	0.999	0.955 1.044	0.96
2	log	1.002	0.963 1.042	0.40	1.001	0.962 1.041	0.97	0.993	0.950 1.038	0.77
3	id	1.015	0.977 1.054	0.30	1.014	0.975 1.054	0.48	1.011	0.967 1.056	0.63
3	log	1.012	0.974 1.052	0.33	1.011	0.973 1.051	0.57	1.004	0.961 1.049	0.86
4	id	1.029	0.992 1.067	0.12	1.029	0.991 1.067	0.14	1.025	0.983 1.068	0.24
4	log	1.020	0.983 1.059	0.23	1.019	0.982 1.058	0.32	1.012	0.970 1.056	0.57
5	id	1.005	0.969 1.041	0.39	1.003	0.967 1.041	0.86	0.999	0.959 1.041	0.96
5	log	1.012	0.976 1.050	0.33	1.011	0.974 1.049	0.56	1.005	0.964 1.047	0.82
CO										
0	id	1.012	0.977 1.049	0.31	1.012	0.977 1.048	0.49	1.011	0.973 1.051	0.57
0	log	1.016	0.962 1.073	0.34	1.015	0.961 1.072	0.60	0.996	0.933 1.063	0.90
1	id	1.004	0.969 1.040	0.39	1.003	0.969 1.039	0.85	1.004	0.966 1.045	0.83
1	log	1.027	0.973 1.083	0.25	1.026	0.972 1.082	0.36	1.010	0.947 1.076	0.77
2	id	1.020	0.984 1.057	0.22	1.020	0.985 1.057	0.27	1.025	0.985 1.067	0.22
2	log	1.024	0.970 1.081	0.28	1.023	0.969 1.080	0.41	1.012	0.948 1.081	0.72
3	id	1.019	0.984 1.055	0.23	1.019	0.984 1.055	0.29	1.024	0.984 1.065	0.25
3	log	1.037	0.984 1.093	0.16	1.037	0.984 1.093	0.18	1.027	0.962 1.096	0.42
4	id	1.029	0.995 1.063	0.10	1.028	0.995 1.063	0.10	1.033	0.995 1.072	0.09
4	log	1.055	1.003 1.110	0.04	1.054	1.003 1.109	0.04	1.047	0.984 1.114	0.15
5	id	0.997	0.965 1.031	0.39	0.997	0.964 1.031	0.85	0.998	0.961 1.037	0.94
5	log	1.014	0.966 1.065	0.34	1.012	0.964 1.063	0.63	1.003	0.944 1.066	0.92

[a] IQR = interquartile range; TR = transformation; id = identical.

237

Table 3. Standard One-Pollutant and Two-Pollutant Models Without Interaction for Best Single-Day Lags for Daily Mortality. Reanalysis of Table 24 in Wichmann et al (2000, p 48)[a]

			Default			Strict			GLM		
	Lag (Days)	TR	RR	CI	P	RR	CI	P	RR	CI	P
One-Pollutant Model											
$NC_{0.01-0.10}$	4	log	1.046	0.997 1.097	0.07	1.045	0.996 1.096	0.07	1.047	0.993 1.103	0.09
$MC_{0.01-2.5}$	0	id	1.031	1.000 1.063	0.05	1.030	0.999 1.062	0.05	1.029	0.993 1.066	0.12
PM_{10}	0	id	1.035	1.001 1.069	0.04	1.034	1.001 1.069	0.04	1.029	0.990 1.069	0.14
SO_2	0	log	1.060	1.011 1.112	0.02	1.061	1.011 1.112	0.02	1.058	1.003 1.116	0.04
NO_2	4	id	1.029	0.992 1.067	0.12	1.029	0.991 1.067	0.14	1.025	0.983 1.068	0.24
CO	4	log	1.055	1.003 1.110	0.04	1.054	1.003 1.109	0.04	1.047	0.984 1.114	0.15
Two-Pollutant Model Without Interaction											
SO_2	0	log	1.078	1.014 1.147	0.02	1.078	1.014 1.147	0.02	1.079	1.010 1.153	0.02
$MC_{0.01-2.5}$	0	id	1.016	0.982 1.052	0.25	1.016	0.982 1.052	0.36	1.014	0.975 1.054	0.49
SO_2	0	log	1.052	0.995 1.112	0.08	1.052	0.995 1.112	0.07	1.052	0.991 1.117	0.10
PM_{10}	0	id	1.023	0.986 1.061	0.20	1.022	0.985 1.060	0.24	1.016	0.975 1.059	0.44
SO_2	0	log	1.064	1.013 1.118	0.02	1.064	1.013 1.118	0.01	1.059	1.002 1.120	0.04
$NC_{0.01-0.1}$	4	log	1.045	0.996 1.097	0.07	1.044	0.995 1.095	0.08	1.045	0.991 1.102	0.10
$NC_{0.01-0.1}$	4	log	1.039	0.986 1.094	0.14	1.038	0.986 1.094	0.16	1.042	0.983 1.103	0.17
$MC_{0.01-2.5}$	0	id	1.034	1.000 1.068	0.05	1.033	1.000 1.068	0.05	1.030	0.992 1.070	0.13
$NC_{0.01-0.1}$	4	log	1.040	0.984 1.100	0.15	1.041	0.985 1.101	0.16	1.048	0.984 1.116	0.15
NO2	4	id	1.007	0.963 1.054	0.38	1.005	0.960 1.052	0.82	0.998	0.947 1.052	0.94
$NC_{0.01-0.1}$	4	log	1.034	0.978 1.093	0.20	1.034	0.978 1.094	0.24	1.045	0.980 1.115	0.18
CO	4	log	1.031	0.970 1.095	0.25	1.029	0.968 1.093	0.37	1.012	0.936 1.094	0.77

[a] TR = transformation; NC = number concentration; MC = mass concentration; id = identical.

Table 4. Two-Pollutant Models with Interaction for Daily Mortality. Reanalysis of Table 25 in Wichmann et al (2000, p 49)[a]

	Lag	TR	RR/IQR Default[b]	Strict	GLM
Immediate Effects Only (lag 0)					
SO_2 at $MC_{0.01-2.5}$ low	0	log	1.078	1.078	1.076
$MC_{0.01-2.5}$ at SO_2 low	0	id	1.016	1.014	0.998
Both			1.095	1.095	1.086
SO_2 at PM_{10} low	0	log	1.051	1.051	1.051
PM_{10} at SO_2 low	0	id	0.994	0.994	0.979
Both			1.068	1.068	1.061
Delayed Effects Only (lag 4)					
$NC_{0.01-0.1}$ at NO_2 low	4	log	1.028	1.029	1.029
NO_2 at $NC_{0.01-0.1}$ low	4	id	0.980	0.976	0.957
Both			1.040	1.039	1.034
$NC_{0.01-0.1}$ at CO low	4	log	1.017	1.017	1.019
CO at $NC_{0.01-0.1}$ low	4	log	1.006	1.002	0.972
Both			1.060	1.058	1.049
Immediate and Delayed Effects (lag 0, lag 4)					
SO_2 at $NC_{0.01-0.1}$ low	0	log	1.047	1.046	1.041
$NC_{0.01-0.1}$ at SO_2 Low	4	log	1.031	1.029	1.030
Both			1.109	1.108	1.104
$MC_{0.01-2.5}$ at $NC_{0.01-0.1}$ low	0	id	1.069	1.068	1.064
$NC_{0.01-0.1}$ at $MC_{0.01-2.5}$ low	4	log	1.071	1.069	1.072
Both			1.101	1.100	1.100

[a] TR = transformation; IQR = interquartile range; MC = mass concentration; NC = number concentration; id = identical.

[b] As published in Wichmann et al (2000).

DISCUSSION

We did not observe major changes in effect estimates when using a GAM with more stringent convergence criteria or when using a fully parametric alternative modeling approach (GLM) for the regressions of daily mortality counts on daily air pollution concentrations in Erfurt, Germany. However, in some cases the effect estimates were slightly smaller using GLM. This could also be due to slight changes introduced into the confounder model by using parametric functions for modeling. The plots (Figure 1) suggest that especially at the ends of the distribution the parametric approach might fit the confounders with less precision than the LOESS functions.

On the other hand, the risk estimates for ultrafine particle NC ($NC_{0.01-0.1}$) in two-pollutant models with various gaseous pollutants even increased a little when using GLM in comparison to GAM and therefore proved to be more stable in these models.

Dominici et al (2002) pointed out that the bias under GAM in S-Plus with default convergence criteria becomes larger when using more degrees of freedom per year to model trend and seasonality. In the initial analyses of this study, only 4 *df*/yr were used to adjust for seasonal variations. Influenza epidemics were included separately, with only few degrees of freedom. This might be the reason that here only little bias from lack of convergence was observed.

The confidence intervals of our risk estimates under GLM are somewhat larger. This might indicate the underestimation of standard errors in the GAM function of S-Plus, which was also observed by Dominici et al (2002). Therefore, some of the confidence intervals for the risk estimates now include the null when estimating the standard errors with GLM. Examples include fine particle mass ($MC_{0.01-2.5}$) and PM_{10}.

Nevertheless, GAM with nonparametric smoothers provides a more flexible approach to adjust for nonlinear confounders in time-series studies of air pollution and health than fully parametric alternatives like GLM.

Reanalysis of the Erfurt data showed that the major interpretations and conclusions of the original study (Wichmann et al 2000) persist even after analyses with more stringent convergence criteria and with alternative modeling strategies.

REFERENCES*

Dominici F, McDermott A, Zeger SL, Samet JM 2002. On the use of generalized additive models in time-series studies of air pollution and health. American Journal of Epidemiology 156: 193–203.

Hastie TJ, Tibshirani RJ 1990. Generalized Additive Models. Monographs on Statistics and Applied Probability 43. Chapman & Hall, London, Glasgow, Weinheim, New York, Tokyo, Melbourne, Madras.

*Bold type identifies publication containing the original analyses revised in this short communication report.

Ibald-Mulli A, Wichmann H-E, Kreyling WG, Peters A 2002. Epidemiological evidence on health effects of ultrafine particles. Journal of Aerosol Medicine 15: 189–201.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak B, Sunyer J, Bacharova L, Schouten JP, Ponka A, Anderson HR 1997. Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. British Medical Journal 314: 1658–1663.

McCullagh P, Nelder JA 1989. Generalized Linear Models. Chapman and Hall, London, 2nd ed.

Oberdörster G, Gelain RM, Ferin J, Weiss B 1995. Association of particulate air pollution and acute mortality: Involvement of ultrafine particles? Inhalation Toxicol 7: 111–124.

Peters A, Wichmann HE, Tuch T, Heinrich J, Heyder J 1997. Respiratory effects are associated with the number of ultra-fine particles. American Journal of Respiratory and Critical Care Medicine 155: 1376–1383.

Pope CA, Dockery DW 1999. Epidemiology of particle effects. In: Air Pollution and Health (Holgate ST, Samet JM, Koren HS, Maynard RL, eds.). Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto, pp. 673–705.

Schwartz J 1994. Air pollution and daily mortality: A review and meta analysis. Environmental Research 64: 36–52.

Schwartz J, Dockery DW, Neas LM 1996. Is daily mortality associated specifically with fine particles? Journal of the Air & Waste Management Association 46: 927–939.

Seaton A, MacNee W, Donaldson K, Godden D 1995. Particulate air pollution and acute health effects. The Lancet 345: 176–178.

von Klot S, Wölke G, Tuch T, Heinrich J, Dockery DW, Schwartz J, Kreyling WG, Wichmann HE, Peters A 2002.

Increased asthma medication use in association with ambient fine and ultrafine particles. European Respiratory Journal 20: 691–702.

Wichmann HE, Peters A 2000. Epidemiological evidence of the effects of ultrafine particle exposure. Philosophical Transactions of the Royal Society of London Series A 358: 2751–2769.

Wichmann H-E, Spix C, Tuch T, Wölke G, Peters A, Heinrich J, Kreyling WG, Heyder J. 2000. Daily Mortality and Fine and Ultrafine Particles in Erfurt, Germany. Part I: Role of Particle Number and Particle Mass. Research Report 98. Health Effects Institute, Cambridge MA.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
CI	confidence interval
CO	carbon monoxide
df	degrees of freedom
GAM	generalized additive model
GLM	generalized linear model
id	identical
IQR	interquartile range
LOESS	locally weighted regression smoothers
MAS	mobile aerosol spectrometer
MC	mass concentration
NC	number concentration
NO_2	nitrogen dioxide
PM	particulate matter
PM_{10}	particulate matter with a diameter less than 10 µm
$PM_{2.5}$	PM with a diameter less than 2.5 µm
RR	relative risk
SO_2	sulfur dioxide
TR	transformation

Airborne Particles and Hospital Admissions for Heart and Lung Disease

Antonella Zanobetti and Joel Schwartz

ABSTRACT

Generalized additive models (GAMs*) are commonly used to control for confounding in studies of the association of air pollution with counts of adverse events. Recently, it was pointed out that the default convergence criterion in most GAM software has been lax and that standard errors were not estimated correctly. Alternative approaches, including natural splines and penalized splines, provide proper standard errors.

In this study we reanalyzed the association between particulate matter (PM) less than 10 µm in aerodynamic diameter (PM_{10}) and hospital admissions for heart and lung disease in 14 US cities, using either GAM software with a strict convergence criterion, natural splines, or penalized splines to control for season and weather. We examined the association with the two-day mean PM_{10} and using constrained and unconstrained distributed lag models. We also reanalyzed our studies of effect modification using metaregression.

We found significant associations between admissions for cardiovascular disease (CVD), chronic obstructive pulmonary disease (COPD), and pneumonia and PM_{10} for the mean of two days and for quadratic and unconstrained distributed lag models. The results are slightly different from those previously published. Because the penalized spline approach has neither the standard error problems of the GAM software nor the sensitivity to knot location of natural splines, we believe those results are preferable.

---

* A list of abbreviations and other terms appears at the end of the section.

INTRODUCTION

GAMs (Hastie and Tibshirani 1990) have been applied in many time-series studies of air pollution and mortality or morbidity. The wide use of GAM in air pollution epidemiology is due to its flexibility in modeling nonlinear factors such as season and weather (Schwartz 1993 and 1994, Katsouyanni et al 2001, Daniels et al 2000).

The GAM function is available in the S-Plus statistical software (Mathsoft, Seattle WA). Estimation in GAM (as written for most current statistical packages) is based on a combination of the local scoring algorithm (to fit the non-Gaussian nature of the data) and the backfitting algorithm.

When the smoothing functions in the linear predictor are parametric then the additive regression model is fit by using weighted least square and the GAM is equivalent to a generalized linear model (GLM). The backfitting algorithm is used within the local scoring algorithm iteration when several nonparametric smoothing functions are used in the model. With only one smooth function, backfitting is not required.

A recent report from Dominici and coworkers (2002) has indicated that the default convergence criteria used in the S-Plus function GAM is relatively lax and may not guarantee convergence. This is not a software problem, merely a reminder that investigators need to pay attention to defaults in statistical software as they are not always appropriate for each problem. Independently, Ramsay and colleagues (2003) reported that the S-Plus GAM function used a shortcut in estimating the covariance of the estimated coefficients that does not properly take into account the correlation between the exposure variables of interest and the smoothed functions of covariates. This may result in biased estimates of the standard errors for the pollution variables.

These two findings have raised questions about the specific results of air pollution time series previously reported and the reliability of the general approach and findings in particular.

We have previously examined the association between particulate air pollution and hospital admission for cardiovascular and respiratory disease in ten US cities

(Zanobetti et al 2000a). In the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) report to the Health Effects Institute (Samet et al 2000), we added four additional cities. In both analyses, the counts of daily hospital admissions for CVD (*International Classification of Diseases, Ninth Edition* [ICD9] 390–429), pneumonia (ICD9 480–487), and COPD (ICD9 490–492 and 494–496) in each city were analyzed in a generalized additive Poisson regression model, using the locally weighted running-line (LOESS; Cleveland and Delvin 1988) as smoothing function. We controlled for temperature, previous day's temperature, relative humidity, seasonality, and barometric pressure. Indicator variables were used to control for day of the week. Air pollution (PM_{10}) was entered linearly and we looked at the mean of PM_{10} on the same and previous day and at distributed lag models.

The results from the 14 cities analysis was used in a subsequent paper (Janssen et al 2002) to evaluate whether differences in prevalence of air conditioning and/or the contribution of different sources to total PM emissions could explain the variability observed in the exposure–effect relations. The data was analyzed using metaregression techniques.

In this paper we report the reanalysis of these previous studies using GAM with LOESS with a stringent convergence criterion, natural spline and penalized spline models. As the NMMAPS report included the 10 US cities from our other paper (Zanobetti et al 2000a; Schwartz 10 Cities this volume; Schwartz et al this volume), we report the results for the 14 US cities.

Natural spline has been applied as an alternative to nonparametric smoothing (Dominici et al 2002), but this model is sensitive to knot locations and knots number. Penalized spline represents an attractive alternative that maintains the flexibility of nonparametric smoothing without the problems associated with GAM. Penalized spline software is advantageous for 3 separate reasons: it does not use backfitting, it lacks sensitivity to knot locations, and it can be used for model building.

METHODS

Penalized spline models were introduced by Eiler and Marx (1996) and have been applied previously in air pollution epidemiology (Zanobetti et al 2000b, Coull et al 2001). The idea of penalized spline is simple. Suppose we want to smooth daily deaths or hospital admissions against temperature and other covariates to obtain flexible estimates of their effects on outcome. We then have a model such as:

$$\text{Cases}_t \sim \text{Poisson}(\exp\{\beta_0 + f(\text{temp}_t) + g(\text{season}_t) + \ldots + \eta PM10\}) \tag{1}$$

We can estimate the smooth function f (and similarly with g, etc) as follows: …

$$f(\text{temp}_t) = \beta_{f1}\text{temp}_t + \beta_{f2}\text{temp}_t^2 + \beta_{f3}\text{temp}_t^3 + \sum_k u_k^f (\text{temp}_t - \kappa_k)_+^3 \tag{2}$$

where K is the number of boundary points (knots) between intervals of temperature, κ_k are the locations of those points, and $(\text{temp}_t - \kappa_k)_+$ is defined as:

$$\begin{cases} \text{temp}_t - \kappa_k, \text{temp}_t > \kappa_k \\ 0, \text{ otherwise} \end{cases}$$

So far, this is simply a piecewise cubic fit or a cubic regressions spline. Splines are polynomial functions that are pieced together to model the association between variables (Wypij 1996). Natural splines are cubic regression spline linear at the extreme intervals (that is, the 2nd and 3rd derivative are zero at the extremes).

If we now fit equation (2) using a penalized likelihood, we constrain changes in parameters u_k at each knot point, effectively reducing the degree of freedom of piecewise fit.

The larger the parameter (called the smoothing parameter) that multiplies the penalty term, the smaller the changes in slope that will minimize this penalized likelihood, and hence the smoother the curve. The degree of smoothness (and hence the value of the parameter) can be estimated by using generalized cross validation.

Software to use penalized splines to fit multiple covariates to a health outcome, including logistic and Poisson regressions, were developed by Wood (2000) and is available in R (The Comprehensive R Archive Network: *http://cran.r-project.org/*). This software uses cubic splines or thin plate splines.

When a penalized regression spline is used, we choose more knots than the expected degree of freedom required. We also use a penalized likelihood to constrain the parameters u_k and therefore reduce the degree of freedom. In the regression spline approach, we choose a number of knots based on the desired degree of freedom and fit the u_k without constraint. This is a less flexible approach that is more likely to be sensitive to knot placement.

We have used both approaches to estimate the effect of PM_{10} on hospital admissions for CVD, COPD, and pneumonia controlling for season, temperature on the day of and day before admission, relative humidity, and

barometric pressure. In addition, we controlled for day of the week (indicator variables).

To define our original models, we chose a separate smoothing parameter in each city and for each outcome to remove seasonality and reduce the residuals of the regression to white noise (ie, remove serial correlation). To allow for city specific divergences, the smoothing parameters for weather variables were also chosen separately in each location, based on the Akaike information criterion.

In this reanalysis we have chosen the degree of freedom to be the same as in our previously published models where we used LOESS function in GAM (Samet et al 2000, Zanobetti et al 2000a) in order to enhance comparability. The autocorrelation of the residuals of the new models was still white noise, and in some cities autoregressive terms were added as in our original models.

To address the question of sensitivity of the effect size estimates to the choice of convergence criteria and maximum iterations in GAM, we refit the originally reported model changing only the convergence criteria. The stricter convergence criteria were the ones recommended by Dominici and coworkers (2002), that is, 10^{-15}.

PM_{10} was analyzed as a linear term for the mean of lag 0 and 1, and we also analyzed quadratic and unconstrained distributed lag models to examine how PM_{10} effects were distributed over different lags. In those models, we included terms for PM_{10} on the same day and each day between one and five days prior to admission to capture the potential for delayed effect of pollution. The unconstrained distributed lag model of order q is:

$$\text{Cases}_t \sim \text{Poisson}(\exp\{\beta_0 + f(\text{temp}_t) + g(\text{season}_t)$$
$$+ \ldots + \eta_0 PM_{10}t + \eta_1 PM_{10}t - 1$$
$$+ \ldots + \eta_q PM_{10}t - q\})$$

The overall impact of a unit change in exposure on one day on admissions over the next q days is given by the sum $\eta_0 + \eta_1 + \ldots + \eta_q$. Even if PM_{10} at time t is correlated with $PM_{10}t - 1 \ldots PM_{10}t - q$, both the η_q and the sum of them will be unbiased estimators of the effects at each lag and of the overall effects. The correlation will make the estimates unstable in a single city, however. Because they are unbiased, combining results across cities will produce more stable unbiased estimates. More detailed description of distributed lag models has been published previously (Zanobetti et al 2000a, Zanobetti et al 2000b, Schwartz 2000).

In the second stage we combined the city specific coefficients $\hat{\eta}_i$, using the multivariate maximum likelihood method of Berkey and coworkers (1998).

We assume that:

$$\hat{\eta}_i \sim N\left(\hat{\eta}, \hat{S}_i + D\right)$$

where $\hat{\eta}_i$ is the vector of $\hat{\eta}$ in city i, $\hat{S}_i$ is the estimated variance–covariance matrix in city i, and D is the random variance–covariance matrix component, reflecting heterogeneity in response among the cities.

RESULTS

MAIN RESULTS

Table 1 shows the combined overall estimate for the three outcomes for the constrained (two-day mean, quadratic distributed lag) and the unconstrained distributed lag model. These show the percentage increase in hospital admissions for a 10 µg/m^3 increase in PM_{10}. These results are for the original GAM models, GAM models with the stricter convergence criteria, natural spline and penalized spline models. The new results generally decrease from those reported in the original GAM models. The penalized spline models results are higher than the natural spline models results and similar to the LOESS with a stringent convergence criterion.

When the analysis was restricted to days when PM_{10} was less than 50 µg/m^3, the results increased compared to the same analysis over all ranges of PM_{10} (< 150 µg/m^3).

We found the main differences for pneumonia. Pneumonia is the outcome measure that shows the greatest seasonality. While pneumonia cases peak in the winter, in many of the cities, there are important peaks of PM_{10} in the summer. It is interesting that when the analysis is restricted to days with PM_{10} below 50 µg/m^3, the effect size doubles. This suggests that the pneumonia results are sensitive to peaks in exposure during a season when pneumonia is barely present. Despite this greater sensitivity of pneumonia results, the general result is that the association of PM_{10} with hospital admissions remains and in most cases is little changed.

EFFECT MODIFICATION

Tables 2 and 3 show the change from baseline PM_{10} effect (as percentage increase in admission per 10-µg/m^3 increase in concentration) associated with a 5-point increase in each measure for models with natural spline and penalized spline respectively.

For COPD and pneumonia, none of the analyzed factors were significant modifiers of the PM_{10} effect estimates in the 14 cities. For CVD we found that the percentage of the population living in poverty, the percentage of the

Table 1. Percentage Increase in Hospital Admission or CVD, COPD, and Pneumonia per 10 µg/m^3 Increase in PM$_{10}$ in 14 US Cities: Combined Effect Using Maximum Likelihood Estimation[a]

	Two-Day Mean		Quadratic Distributed Lag		Unconstrained Distributed Lag		Two-Day Mean PM < 50	
	%	95% CI	%	95% CI	%	95% CI	%	95% CI
COPD								
Original GAM	1.81	1.16 2.47	2.94	0.21 5.74	2.88	0.22 5.61	2.60	1.40 3.81
GAM with convergence criterion 10^{-15}	1.71	0.95 2.48	2.53	1.20 3.88	2.53	1.21 3.87		
Natural spline	1.32	0.56 2.08	2.10	0.63 3.59	2.11	0.63 3.61	2.21	1.02 3.41
Penalized spline	1.56	0.86 2.28	2.39	1.13 3.67	2.54	1.20 3.89		
CVD								
Original GAM	1.17	1.01 1.33	1.05	0.67 1.43	1.06	0.67 1.46	1.45	1.12 1.78
GAM with convergence criterion 10^{-15}	0.99	0.79 1.19	1.09	0.81 1.38	1.12	0.84 1.40		
Natural spline	0.96	0.71 1.20	1.06	0.72 1.40	1.11	0.79 1.44	1.32	0.77 1.87
Penalized spline	1.00	0.80 1.19	1.09	0.81 1.37	1.12	0.83 1.41		
Pneumonia								
Original GAM	1.90	1.44 2.37	1.87	0.72 3.04	2.07	0.93 3.23	2.46	1.16 3.78
GAM with convergence criterion 10^{-15}	1.71	1.16 2.26	1.47	0.86 2.09	1.62	0.95 2.29		
Natural spline	0.57	0.04 1.10	−0.02	−0.65 0.61	0.01	−0.64 0.67	1.06	0.06 2.07
Penalized spline	1.23	0.49 1.98	0.64	0.03 1.25	0.80	0.14 1.46		

[a] CVD = cardiovascular disease; COPD = chronic obstructive pulmonary disease.

Table 2. Effect Modification by Percentages of the Population That Are College Educated, Unemployed, Living in Poverty or Nonwhite. Results Are Shown for a 10 µg/m^3 Increase in PM$_{10}$ and a 5% Point Increase in 14 Cities. Models with Natural Spline

	College Educated		Unemployed		Poverty		Nonwhite	
	%	95% CI	%	95% CI	%	95% CI	%	95% CI
CVD[a]	−0.17	−0.42 0.07	0.78	0.42 1.14	0.46	0.12 0.79	0.10	0.01 0.19
COPD[a]	0.60	−0.47 1.67	−0.14	−2.60 2.38	−0.96	−2.65 0.76	−0.22	−0.65 0.20
Pneumonia	−0.03	−0.52 0.47	0.37	−0.70 1.45	0.46	−0.32 1.26	0.16	−0.03 0.34

[a] CVD = cardiovascular disease; COPD = chronic obstructive pulmonary disease.

population that was nonwhite, and the percentage of unemployed were modifiers of hospitalizations for cardiovascular disease. This finding held for both types of models, even though with natural spline models the results were stronger. The effect of PM$_{10}$ was greater in communities with higher levels of these indicators. This differs from our previous results, where they were not significant modifiers, and suggests that social deprivation may convey susceptibility to the effects of particles. This could be through increased exposure (eg, because of less air conditioning), increased prevalence of predisposing diseases (eg, diabetes mellitus), or other factors.

In the metaregression with hospitalization rates, we found that the coefficients for hospital admission rates for CVD, COPD, or pneumonia were not associated with modification of the PM$_{10}$ effect estimates.

Table 3. Effect Modification by Percentages of the Population That Are College Educated, Unemployed, Living in Poverty or Nonwhite. Results Are Shown for a 10 mg/m^3 Increase in PM$_{10}$ and a 5% Point Increase in Effect Modifiers in 14 Cities. Models with Penalized Spline

	College Educated		Unemployed		Poverty		Nonwhite	
	%	95% CI	%	95% CI	%	95% CI	%	95% CI
CVD[a]	−0.05	−0.28 0.17	0.45	0.01 0.88	0.26	−0.08 0.60	0.08	0.00 0.16
COPD[a]	0.87	0.00 1.76	0.13	−2.23 2.55	−1.48	−3.14 0.21	−0.17	−0.62 0.29
Pneumonia	0.17	−0.31 0.64	0.24	−0.89 1.39	0.19	−0.67 1.06	0.08	−0.13 0.29

[a] CVD = cardiovascular disease; COPD = chronic obstructive pulmonary disease.

When examining the correlation of PM$_{10}$ with temperature and relative humidity, we found the same results as previously. In the metaregression, we found that the coefficient for temperature and relative humidity were not significant for all the three outcomes.

In our previous studies we did not find confounding due to other pollutants. These results are confirmed in this reanalysis by the metaregression estimates, shown in Figure 1 for natural spline models and Figure 2 for penalized spline models. In these plots the baseline estimate is the result of the distributed lag metaanalysis. Plotted above each pollutant is the estimated intercept in the metaregression of the PM$_{10}$ coefficients against the slopes between that copollutant and PM$_{10}$. These plots show little evidence of PM$_{10}$ effect confounded by other pollutants.

AIR CONDITIONING AND SOURCE-RELATED VARIABLES

Table 4 shows the results for the metaregression between the coefficients in the 14 cities and homes with air conditioning using the iterative maximum likelihood approach (Berkey et al 1995). The table presents the results for the original GAM models (Janssen et al 2002) and the coefficients from the natural spline and penalized spline models. For all cities, central air conditioning was not associated with PM$_{10}$ coefficients. In the model that adjusted for whether the cities were characterized by winter peaking PM$_{10}$ concentrations (yes/no), however, the coefficients for CVD-related hospital admissions decreased significantly with increasing percentage of central air conditioning.

Table 5 shows the results of the regression analysis of the source-related variables. All the results were adjusted for central air conditioning percentage. With all models the coefficients for hospital admissions for CVD increased significantly with increasing percentage of PM$_{10}$ from highway vehicles, highway diesels, and oil combustion.

The results found previously for air conditioning and source-related variables still hold while using natural spline and penalized spline models.

Figure 1. Metaregression adjustment for copollutants. Combined estimated effects of PM$_{10}$ on cardiovascular disease, chronic obstructive pulmonary disease (COPD), and pneumonia without (base) or with adjustment for individual gaseous pollutants. Natural spline models.
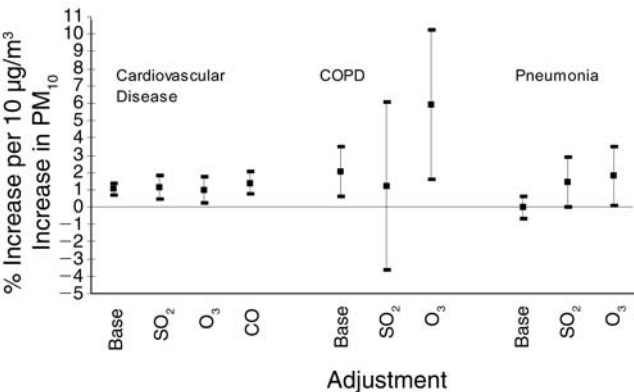

Figure 2. Metaregression adjustment for copollutants. Combined estimated effects of PM$_{10}$ on cardiovascular disease, chronic obstructive pulmonary disease (COPD), and pneumonia without (base) or with adjustment for individual gaseous pollutants. Penalized spline models.
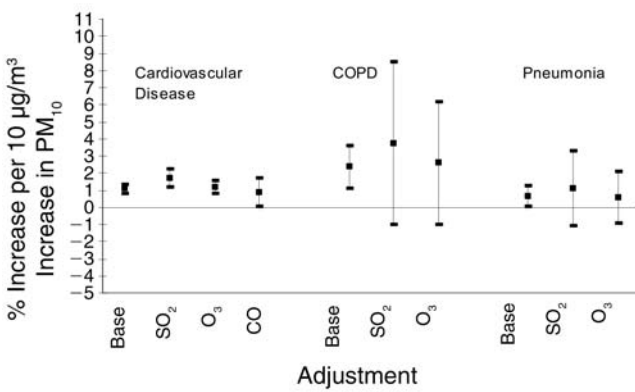
245

Table 4. Percentage Change in the Coefficient of the Effect of Ambient PM_{10} on Hospital Admissions for CVD[a] for an IQR[b] Increase in the Percentage of Homes With Central Air Conditioning. Results for the Original GAM Models With LOESS, Natural Spline and Penalized Spline Models

	n	IQR	Original GAM			Natural Spline			Penalized spline		
			% Change	SE (%)	t	% Change	SE (%)	t	% Change	SE (%)	t
All cities	14	20.44	−15.2	14.8	1.03	−13.55	14.88	0.91	−11.97	14.11	0.85
Nonwinter peaking cities	9	20.44	−50.3	17.4	2.89	−44.11	20.15	2.19	−38.41	17.79	2.16
Winter peaking cities	5	20.44	−51.7	13.8	3.75	−6.10	40.27	0.15	−41.47	39.63	1.05
Adjusted for winter peaking cities	14	20.44	−50.5	14.6	3.46	−37.65	17.10	2.20	−38.70	16.14	2.40

[a] CVD = cardiovascular disease.

[b] IQR = interquartile range.

Table 5. Percentage Change in the Coefficient of the Effect of Ambient PM_{10} on Hospital Admissions for CVD[a] for an IQR[b] Increase in the Percentage of PM_{10} from Different Sources, Population Density, and VMT[a]/$mile^2$, Adjusted for Central Air Conditioning (%). Results for the Original GAM with LOESS, Natural Spline and Penalized Spline Models

	IQR	Original GAM Results			Natural Spline			Penalized Spline		
		% Change	SE (%)		% Change	SE (%)		% Change	SE (%)	
Sources of PM_{10}										
Highway vehicles	2.64	58	9.9	5.86	51.13	14.67	3.49	35.09	14.28	2.46
Highway diesels	1.24	55.6	9.4	5.91	50.06	14.15	3.54	34.98	13.80	2.53
Coal combustion	0.29	0.6	2.6	0.23	1.27	2.89	0.44	0.38	2.70	0.14
Oil combustion	0.29	37.5	9.3	4.03	31.25	11.55	2.71	35.14	11.06	3.18
Wood burning	1.37	2.7	3.2	0.84	5.42	3.00	1.81	5.55	2.68	2.07
Metal processing	4.72	29	13	2.23	18.40	16.36	1.12	9.62	15.61	0.62
Fugitive dust	21.53	−49.4	16.5	2.99	−53.06	16.58	3.20	−27.90	18.89	1.48
Population density	1.60	22.4	7.8	2.87	13.47	10.83	1.24	11.69	9.22	1.27
VMT/$mile^2$	13.07	21.2	7.4	2.86	12.67	10.39	1.22	10.84	8.94	1.21

[a] CVD = cardiovascular disease; VMT = vehicle miles traveled.

[b] IQR = interquartile range.

DISCUSSION

In this paper we present the results of the association between PM_{10} and hospital admissions for heart and lung disease in 14 US cities, using natural spline and penalized splines to control for season and weather. We also applied GAM with LOESS but with a stringent convergence criterion. We found a significant association with all three outcomes. The order of magnitude of the results is greater for LOESS, almost the same with penalized spline and a little smaller with natural spline.

For admissions from heart disease and COPD, these results are only slightly different from the results previously published using GAM models.

The pneumonia results, in contrast, differed substantially from those reported earlier using GAM models. That difference was lessened by excluding high pollution days. This suggests that the pneumonia results are sensitive to peaks in exposure during a season when pneumonia is barely present. Despite this greater sensitivity of pneumonia results, the general result is that the association of PM_{10} with hospital admissions remains and in most cases is little changed.

We also presented the reanalysis for air conditioning and source-related particles as modifiers of the PM_{10} effect on hospital admissions. We still found a decreased PM_{10} effect with increasing percentage of homes with central air conditioning. The results also hold true for source-related

particles, with higher effects for PM_{10} from increasing percentage of PM_{10} from highway vehicles, highway diesels, and oil combustion.

Penalized spline models represent an attractive approach to modeling the potentially nonlinear dependence of a health outcome on several predictors. Conceptually, it bridges the gap between regression splines, which are straightforward, and smoothing splines, which are more flexible, less sensitive, but more complex. Moreover they eliminate all of the problems identified with GAM models: weak convergence criteria, unbiased estimates of linear coefficients (eg, for air pollution), and their standard errors.

Penalized splines also offer a more generalized framework for analyzing a wide variety of study designs. For example, Coull and coworkers (2000) analyzed a longitudinal panel study, incorporating random intercepts and random slopes for pollution, while controlling for a potentially nonlinear effect of temperature and season. Using a two-stage approach, survival analysis can also be conducted in a generalized linear model framework (Thurston et al 2002), and hence the generalized linear mixed model allows extension to penalized splines for these questions as well.

REFERENCES*

Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA.1998. Meta-analysis of multiple outcomes by regression with random-effects. Stat Med 17: 2537–2550.

Cleveland WS and Devlin SJ. 1988. Robust locally-weighted regression and smoothing scatter plots. J Am Stat Assoc 74:829–36.

Coull BA, Schwartz J, Wand MP. 2001. Respiratory health and air pollution: Additive Mixed Model Analysis. Biostatistics 2: 337–49.

Daniels MJ, Dominici F, Samet JM, Zeger SL. 2000. Estimating particulate matter mortality dose response curves and threshold levels: an analysis of daily time series for the 20 largest US cities. Am J Epidemiol 152:397–406.

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On generalized additive models in time series studies of air pollution and health. Am J Epidemiol 156: 193–203.

Eilers PHC, Marx BD. 1996. Flexible smoothing with B-splines and penalties (with discussion). Statistical Science 11:89–121.

Hastie TJ and Tibshirani RJ. 1990. Generalized additive models. Chapman and Hall. New York, NY.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, LeTertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiology 125: 521–31.

Janssen NAH, Schwartz J, Zanobetti A, Suh HH. 2002. Air conditioning and source-specific particles as modifiers of the effect of PM_{10} on hospital admissions for heart and lung disease. Environ Health Perspect 110:43–49.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwartz J, Zanobetti A. 2000. The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Schwartz J. 1993. Air pollution and daily mortality in Birmingham, Alabama. Am J Epidemiol 137:1136–1147.

Schwartz J. 1994. Nonparametric smoothing in the analysis of air pollution and respiratory illness. Canadian J Stat 22: 471–487.

Schwartz J. 2000. The distributed lag between air pollution and daily deaths. Epidemiology 11: 320–326.

Thurston SW, Eisen EA, Schwartz J. 2002. Smoothing in survival models: an application to workers exposed to metalworking fluids. Epidemiology 13: 685–692.

Wood, SN. 2000. Modeling and smoothing parameter estimation with multiple quadratic penalties. JRSS B 62: 413–428.

Wypij D. 1996. Spline and smoothing approaches to fitting flexible models for the analysis of pulmonary function data. American Journal of Respiratory & Critical Care Medicine 154: S223–S228.

Zanobetti A, Schwartz J, Dockery DW. 2000a. Airborne particles are a risk factor for hospital admissions for heart and lung disease. Environ Health Perspect 108:1071–1077.

Zanobetti A, Wand MP, Schwartz J, and Ryan L. 2000b. Generalized Additive Distributed Lag Models: Quantifying Mortality Displacement. Biostatistics 1: 279–292.

* Bold type identifies publications containing the original analyses revised in this short communication report.

ABBREVIATIONS AND OTHER TERMS

CI	confidence interval
CO	carbon monoxide
COPD	chronic obstructive pulmonary disease
CVD	cardiovascular disease
GAM	generalized additive model
GLM	generalized linear model
ICD9	*International Classification of Diseases, Ninth Edition*
LOESS	locally weighted running-line
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
O_3	ozone
PM	particulate matter
PM_{10}	particulate matter less than 10 µm in aerodynamic diameter
SO_2	sulfur dioxide

Multicity Assessment of Mortality Displacement Within the APHEA2 Project

Antonella Zanobetti and Joel Schwartz

ABSTRACT

Generalized additive models (GAMs*) are commonly used to control for confounding in studies of the association of air pollution with counts of adverse events. Recently, it was pointed out that the default convergence criterion in most GAM software has been lax and that standard errors were not estimated correctly. Alternative approaches, including natural splines and penalized splines provide proper standard errors.

We previously examined the mortality displacement issue by means of a distributed lag model, assessing the effect of exposure for up to 40 subsequent days in a multicity hierarchical modeling approach within the Air Pollution and Health: a European Approach (APHEA2) study. If deaths due to air pollution were only being brought forward by a few days, one would expect to see negative associations between air pollution today and the number of deaths occurring in the future, from which today's deaths were borrowed. Instead, we found evidence for delayed addition effects of pollution. This reanalysis readdresses that issue two ways: using GAM with a stringent convergence criterion and using penalized spline models. We continue to find no evidence for harvesting and some evidence for additional delayed effects of pollution.

INTRODUCTION

One important issue in environmental epidemiology is the harvesting phenomenon or short-term mortality displacement. The public health significance of short-term effects of air pollution, in fact, varies depending on the extent of time the deaths are being advanced. Assume that at any given time there is a pool of people at elevated risk of dying because of acute or chronic conditions. Then an air pollution episode could increase the risk for those individuals, increasing the death rate out of the pool and decreasing the pool size. On subsequent days, the number of deaths observed will be less than expected, inducing a negative association with air pollution at those lags. The finite nature of the risk pool therefore creates the possibility of this negative association, with the result that many of the early deaths may only be advanced by a short period. This is referred to as *mortality displacement* and several papers have addressed this issue in single-city studies (Zeger et al 1999, Schwartz 2000a, 2001, Zanobetti et al 2000).

In a previously published paper (Zanobetti et al 2002), we examined the mortality displacement issue with a distributed lag model in a multicity hierarchical modeling approach within the APHEA2 study. We chose the ten largest cities in the APHEA2 study; in the first stage of the analysis, we fit a Poisson regression model and a polynomial distributed lag model with up to 40 days of delay in each city. If the deaths following today's pollution exposure are borrowed from, for example, next week, then we would expect a negative correlation between today's exposure and next week's death count. The distributed lag model allows us to directly examine this question. We controlled for seasonal patterns, weather, influenza epidemics, holidays, and day of the week. The models were built following the APHEA2 method protocol (Katsouyanni et al 2001). In the second stage we combined the city-specific results using the multivariate maximum likelihood method (Berkey et al 1998), and we found that the size of the effect estimate for airborne particles more than doubled when we considered longer term effects.

To control for season and weather (temperature and relative humidity), we used nonparametric smooth functions.

Recently several authors have pointed out that there are problems with implementation of GAMs in current statistical packages and use of nonparametric smoothing function (Dominici et al 2002, Ramsay et al 2003). These two findings have raised questions about the specific results of

---

air pollution time series previously reported and about the reliability of the general approach and findings in particular.

In this paper we report the reanalysis of this study using GAM with a locally weighted smoother (LOESS) with a stringent convergence criterion and penalized spline models.

Because the penalized spline models neither use back-fitting nor misestimate standard errors, they represent an alternative approach to control for nonlinear dependence on season and weather.

METHODS

ORIGINAL ANALYSIS

For each city the original analysis fit a generalized additive Poisson regression, modeling the logarithm of the expected value of daily deaths as a sum of smooth functions of the predictor variables. We controlled for seasonal patterns, long-term time trends for weather, influenza epidemics, holidays, and day of the week. The models were built following the APHEA2 method (Katsouyanni et al 2001). All models controlled for temperature and humidity on the same day using nonparametric smooth functions (Cleveland and Delvin 1988). In addition, we examined whether nonparametric functions of weather variables on the previous day or up to three previous days or the average of a few days improved model fit. We similarly chose the degree of freedom for each weather variable to minimize the Akaike information criterion (AIC; Akaike 1973).

To control for seasonality, we chose a separate smoothing parameter in each city, and the parameters were chosen so as to reduce the residuals of the regression to white noise. In some cities, it was necessary to introduce autoregressive terms.

The goal of our original analysis was to estimate the dependence of daily deaths (on day t) on particulate matter less than 10 μm in diameter (PM_{10}) on that day and up to the previous 40 days. If the pollution-related deaths were only being advanced by a few days to a few weeks, we would see this effect as a negative association between air pollution and deaths several days to weeks subsequently. The net effect of air pollution, net of any such short-term rebound up to 40 days, is the sum of the effect estimates for all 41 days.

The unconstrained distributed lag model of order q is:

$$\text{Cases}_t \sim \text{Poisson}\left(\exp\{\alpha + \text{covariates} + \beta_0 PM_{10\,t}\right.$$
$$\left. + \beta_1 PM_{10\,t-1} + ... + \beta_q PM_{10\,t-q}\}\right)$$

Because this model produces unstable estimates for large q, an alternative is to provide more stability constraining the coefficients to vary smoothly with lag number. A polynomial distributed lag constrains the β_j to follow a polynomial pattern in the lag number, that is:

$$\beta_j = \sum_{k=0}^{d} \eta_k j^k, \text{ for } j = 0...q$$

where j is the number of lag of delay and k is the degree of the polynomial.

More detailed description of distributed lag models has been published (Zanobetti et al 2000a, Zanobetti et al 2000b, Schwartz 2000b).

In the second stage we combined the city specific coefficients, $\hat{\beta}_j$, using the multivariate maximum likelihood method (Berkey et al 1998).

We assume that

$$\hat{\beta}_j \sim N(\hat{\beta}, \hat{S}_j + D)$$

where $\hat{\beta}_j$ is the vector of $\hat{\beta}$ in city j, $\hat{S}_j$ is the estimated variance–covariance matrix in city j, and D is the random variance–covariance matrix component, reflecting heterogeneity in response among the cities.

REANALYSIS APPROACH

In this reanalysis we have chosen the degree of freedom to be the same as in our previously published models where we used the LOESS function in GAM (Zanobetti et al 2002) in order to enhance comparability.

To address the question of sensitivity of the effect size estimates to the choice of convergence criteria and maximum iterations in GAM, we refit the originally reported model changing only the convergence criteria. The stricter convergence criteria were the ones recommended by Dominici and coworkers (Dominici et al 2002), that is 10^{-15}.

In the second part of reanalyses, we used penalized spline models with the same degree of freedom for the covariates as in the original models.

Penalized spline models (Eiler and Marx, 1998) are cubic regressions spline fit using a penalized likelihood. For example, suppose we want to smooth daily deaths or hospital admissions against temperature and other covariates to obtain flexible estimates of their effects on outcome. We then have a model such as:

$$\text{Cases}_t \sim \text{Poisson}(\exp\{\alpha$$
$$+ f(\text{temp}_t) + g(\text{season}_t) \tag{1}$$
$$+ ... + \beta_0 PM_{10}\})$$

We can estimate the smooth function f (and similarly with g, etc) as follows:

$$f(\text{temp}_t) = \beta_{f1}\text{temp}_t + \beta_{f2}\text{temp}_t^2 \quad (2)$$
$$+ \beta_{f3}\text{temp}_t^3 + \sum_k u_k^f (\text{temp}_t - \kappa_k)_+^3$$

where K is the number of boundary points (knots) between intervals of temperature, κ_k are the locations of those points, and $(\text{temp} - \kappa_k)_+$ is defined as:

$$\begin{cases} \text{temp}_t - \kappa_k, & \text{temp}_t > \kappa_k \\ 0, & \text{otherwise} \end{cases}$$

So far, this is simply a cubic regressions spline. If we now fit equation (2) using a penalized likelihood, we constrain changes in parameters u_k at each knot point, effectively reducing the degree of freedom of the piecewise fit.

The larger the parameter (called the smoothing parameter) that multiplies the penalty term, the smaller the changes in slope that will minimize this penalized likelihood and hence the smoother the curve. The degree of smoothness (and hence the value of the parameter) can be estimated by using generalized cross validation.

When a penalized regression spline is used, we choose more knots than the expected degrees of freedom required, and we use a penalized likelihood to constrain the u_k and therefore reduce the degree of freedom.

Software to use penalized splines to fit multiple covariates to a health outcome, including logistic and Poisson regressions, has been developed by Wood (2000) and is available in R (The Comprehensive R Archive Network: *http://cran.r-project.org/*). This software uses cubic splines or thin plate splines.

RESULTS

Table 1 shows the ten cities, their populations, and the means and standard deviations of the number of daily deaths and environmental variables. Other details of the baseline models for each city have been published previously (Katsouyanni et al 2001).

Tables 2 and 3 show city specific and combined results for the original GAM models, GAM models with the convergence criterion 10^{-15}, and penalized spline models. Table 2 presents the estimated regression coefficients of PM_{10} (per 10 $\mu g/m^3$, standard errors and t values) for the mean of the current and previous day of PM_{10} and Table 3 for the 4th degree polynomial distributed lag model.

The effect size estimates were slightly reduced using the stricter convergence criteria, and slightly more using the penalized spline model. The reductions are lower for the 2-day mean of PM_{10} and higher for the distributed lag model with 40 lags of delay, but the effect estimates still remain considerably higher (more than double) than estimates using a mean of two days exposure.

Overall the changes were small and leave the general impression that there is no consistent difference between the new results and the previously published results.

Table 1. Population, Mean and Standard Deviation of Daily Deaths and Environmental Variables in the 10 Cities

	Population	Total Mortality		PM_{10} ($\mu g/m^3$)		Temperature		Humidity	
	($\times$ 1000)	(mean)	(SD)	(mean)	(SD)	(mean)	(SD)	(mean)	(SD)
Athens	3073	72.9	13.2	42.7	12.9	17.8	7.4	61.7	13.6
Budapest	1931	80.0	11.6	41.0	9.1	12.8	8.8	70.1	12.6
Lodz	828	29.5	6.3	53.5	15.5	8.4	8.4	79.0	12.4
London	6905	168.5	25.2	28.8	13.7	11.8	5.4	69.3	11.3
Madrid	3012	60.8	11.1	37.8	17.7	14.5	7.4	61.8	16.7
Paris	6700	123.3	15.7	22.5	11.5	12.1	6.5	75.6	12.5
Prague	1212	38.2	7.2	76.2	45.7	11.0	8.0	69.4	14.1
Roma	2775	56.2	10.4	58.7	17.4	16.8	6.7	61.6	11.9
Stockholm	1126	28.9	6.1	15.5	7.9	7.7	8.1	71.4	15.8
Tel Aviv	1141	27.4	6.3	50.3	57.5	20.6	5.4	65.6	11.0

Table 2. APHEA2: Total Mortality and PM_{10} in 10 European Cities. City-Specific and Combined Results for the Estimated PM_{10} Effect ($\times$ 1000) for the Mean PM_{10} Lags 0–1

| | PM₁₀ Mean Lags 0–1 | | | | | | | | |
| | Original GAM | | | GAM with Convergence Criterion 10^{-15} | | | Penalized Spline Model | | |
	β	SE	t	β	SE	t	β	SE	t
Athens	1.64	0.29	5.60	1.58	0.29	5.41	1.41	0.30	4.68
Budapest	0.28	0.46	0.61	0.13	0.46	0.28	−0.12	0.47	−0.25
Lodz	0.59	0.42	1.41	0.58	0.42	1.38	0.49	0.47	1.03
London	0.70	0.18	3.94	0.69	0.18	3.88	0.57	0.18	3.22
Madrid	0.52	0.24	2.22	0.49	0.24	2.10	0.41	0.25	1.65
Paris	0.42	0.23	1.82	0.39	0.23	1.71	0.38	0.26	1.49
Prague	0.11	0.18	0.60	0.11	0.18	0.57	0.06	0.19	0.33
Rome	1.51	0.27	5.56	1.51	0.27	5.56	1.28	0.28	4.60
Stockholm	0.36	0.88	0.41	0.36	0.88	0.40	0.63	0.86	0.74
Tel Aviv	0.67	0.26	2.62	0.61	0.26	2.37	0.49	0.27	1.81
Metaanalysis	0.70	0.14	5.13	0.67	0.14	4.80	0.57	0.15	3.82

Table 3. Total Mortality and PM_{10} in 10 European Cities. City-Specific and Combined Results for the Estimated PM_{10} Effect ($\times$ 1000) for the 4th Degree Distributed Lag Models with 40 Lags

| | 4th Degree Distributed Lag Models | | | | | | | | |
| | Original GAM | | | GAM with Convergence Criterion 10^{-15} | | | Penalized Spline Model | | |
	β	SE	t	β	SE	t	β	SE	t
Athens	3.54	0.57	6.16	3.63	0.57	6.32	3.38	0.91	3.73
Budapest	1.41	0.86	1.65	0.43	0.85	0.50	0.36	1.09	0.33
Lodz	3.88	0.62	6.30	3.80	0.62	6.16	3.72	0.94	3.96
London	1.17	0.44	2.63	1.16	0.44	2.61	0.99	0.52	1.88
Madrid	2.34	0.52	4.52	1.87	0.52	3.61	0.38	0.82	0.46
Paris	2.54	0.46	5.53	2.31	0.46	5.03	1.86	0.68	2.72
Prague	0.72	0.34	2.13	0.61	0.34	1.82	0.03	0.51	0.06
Rome	−0.74	0.48	−1.55	−0.72	0.48	−1.52	−0.62	0.74	−0.84
Stockholm	1.93	2.02	0.95	1.93	2.02	0.95	1.82	2.08	0.87
Tel Aviv	0.65	0.38	1.71	0.59	0.38	1.56	0.47	0.42	1.11
Metaanalysis	1.61	0.30	5.32	1.45	0.30	4.79	1.08	0.40	2.73

DISCUSSION

In this paper we present the results of reanalyses of the mortality displacement issue analyzed with a distributed lag model in a multicity hierarchical modeling approach, within APHEA2 study. The reanalysis was done using GAM with a stringent convergence criterion and with penalized splines models.

The results are slightly different from the previously published. Our study confirms that the effects observed in daily time series studies are not due primarily to short-term mortality displacement. The effect size estimate for airborne particles more than doubles when we considered longer term effects, which has important implications for risk assessment.

This analysis, showing qualitatively unchanged results compared to the previously published study, can be added to other reported reanalyses of time series studies to provide confidence that the overall conclusions previously reached about this extensive literature should remain unchanged.

REFERENCES*

Akaike H. 1973. Information Theory and an extension of the maximum likelihood principal. In: BN Petrov and F Csaki eds: 2nd International Symposium on Information Theory, Akademiai Kaiado, Budapest.

Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA.1998. Meta-analysis of multiple outcomes by regression with random-effects. Stat Med 17: 2537–2550.

Cleveland WS and Devlin SJ. 1988. Robust locally-weighted regression and smoothing scatterplots. J Am Stat Assoc 74:829–36.

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On generalized additive models in time series studies of air pollution and health. Am J Epidemiol 156: 193–203.

Eilers PHC, Marx BD. 1996. Flexible smoothing with B-splines and penalties (with discussion). Statistical Science 11:89–121.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, LeTertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braun-

stein R, Pekkanen J,Schindler C, Schwartz J. 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiol 125: 521–31.

Murray CJ, Nelson CR. 2000. State-space modeling of the relationship between air quality and mortality. J Air Waste Manag Assoc 50: 1075–80.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Schwartz J. 2000a. Harvesting and long term exposure effects in the relation between air pollution and mortality. Am J Epidemiology 151: 440–448.

Schwartz J. 2000b. The distributed lag between air pollution and daily deaths. Epidemiology 11:320–326.

Schwartz J. 2001. Is There Harvesting in the Association of Airborne Particles with Daily Deaths and Hospital Admissions? Epidemiol 12:55–61.

Wood, SN. 2000. Modeling and smoothing parameter estimation with multiple quadratic penalties. JRSS B 62: 413–428.

Zanobetti A, Wand MP, Schwartz J, and Ryan L. 2000. Generalized Additive Distributed Lag Models: Quantifying Mortality Displacement. Biostatistics 1,3:279–292.

Zanobetti A, Schwartz J, Samoli E, Gryparis A, Touloumi G, Atkinson R, Le Tertre A, Bobros J, Celko M, Goren A, Forsberg B, Michelozzi P, Rabczenko D, Ruiz EA, Katsouyanni K. 2002. The temporal pattern of mortality responses to air pollution. Epidemiology 13:87–93.

Zeger SL, Dominici F, Samet J. 1999. Harvesting-resistant estimates of air pollution effects on mortality. Epidemiology 10:171–175.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
APHEA2	Air Pollution and Health: a European Approach
GAM	generalized additive model
LOESS	locally weighted smoother
PM$_{10}$	particulate matter less than 10 μm in diameter

---

* Bold type identifies publication containing the original analyses revised in this short communication report.

Commentary on Revised Analyses of Selected Studies

A Special Panel of the Health Review Committee

INTRODUCTION

In May 2002, while conducting exploratory sensitivity analyses of the National Morbidity, Mortality, and Air Pollution Study (NMMAPS*) data, Francesca Dominici and colleagues (2002) from Johns Hopkins University identified unexpected findings associated with applying generalized additive models (GAMs) to time-series data. The team subsequently determined that programming in the GAM function of S-Plus had default settings that allowed the iterative process used in estimating the effect (regression coefficients) to terminate prematurely. This characteristic led to biased estimates of effect. Independently, investigators at Health Canada determined that a programming shortcut used in calculating the standard errors for the regression coefficients underestimated the true standard errors under certain circumstances (Ramsay et al 2003).

The NMMAPS investigators (including both the Johns Hopkins and Harvard teams) promptly conducted new analyses of their data and submitted reports to HEI. An HEI Special Panel evaluated these reports and prepared the separate Commentary published elsewhere in this Special Report.

After the Johns Hopkins investigators alerted the scientific community, some investigators of other time-series studies who had employed GAMs proceeded to conduct new analyses. Many of these studies originally had been cited in the draft document of the US Environmental Agency (EPA) 2002 Air Quality Criteria for Particulate Matter (PM) and were likely to influence the EPA process for assessing the current National Ambient Air Quality Standards for PM. Thus the EPA deemed it necessary to revise the analyses of these

studies to update results. The EPA identified those key studies involving GAMs and cited in the Criteria Document and requested that the investigators carry out revised analyses of their data. The specific request was twofold: (1) to reanalyze the original data using the same nonparametric approach that was used originally, but with stricter convergence criteria and maximum iterations, and (2) to examine and report on the sensitivity of results to nonparametric smoothing models compared with fully parametric models. The use of fully parametric models also aimed to obtain more accurate standard errors. HEI was asked to constitute an HEI Special Panel to review the resulting short communication reports and to prepare this Commentary on these reports.

Nineteen primary authors submitted 21 short communication reports presenting results from analyses originally reported in 37 published original articles and reports. In addition to submitting their reports to HEI in a timely manner, all authors were most helpful in providing additional information to the Panel throughout the review process.

In this Commentary, the HEI Special Panel presents a tabular summary of the principal findings of each revised analysis and a discussion of the revised findings. A detailed discussion of the revised analyses of NMMAPS II data is included in the Commentary prepared by the HEI NMMAPS Special Panel. The present discussion focuses on the new results of other studies: the impact of changing the default convergence criteria and allowing sufficient iterations in the GAM software to reach convergence. The subsequent discussion addresses the impact of analytic models that calculate more accurate standard errors of the estimates of effect. The Review Panel then assesses the impact of conducting sensitivity analyses even though these additional analyses were not requested by EPA. Finally, we offer the Panel's overall conclusions from the revised analyses and their interpretation of findings from observational studies of short-term air pollution.

---

* A list of abbreviations and other terms appears at the end of the Commentary.

This document has not been reviewed by public or private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views of these parties, and no endorsements by them should be inferred.

RESULTS

The revised analyses conducted at the request of EPA sought to evaluate two factors: The first was the sensitivity of effect estimates to the choice of convergence criteria and maximum iterations in GAM. The second was the sensitivity to models that calculate more accurate standard errors, including generalized linear models (GLMs) with natural cubic splines (GLM-natural splines) and other methods selected by the investigators. EPA guidelines to authors suggested fitting a parametric model to the data with the same degrees of freedom (*df*) as for the original nonparametric model. "The specification of the parametric model should be chosen to most nearly approximate the degree of control of time-varying potential confounders afforded by the original GAM-based analysis" (L Grant, electronic communication to investigators and HEI, September 2002). Because of time constraints, authors were encouraged to submit results of additional sensitivity analyses to peer-reviewed journals, rather than include them within the short communication reports.

Commentary Tables 1–3 categorize results according to overall qualitative changes in estimates of effect presented in the short communication reports, except those relating to NMMAPS data, which are discussed in the Commentary on Revised Analyses of NMMAPS II Data. Results from GAM with stricter convergence criteria and maximum iterations (GAM-strict) were compared to GAM with default convergence criteria (GAM-default) (Commentary Table 1); GLM-natural splines were compared to GAM-strict (Commentary Table 2); and estimates obtained by varying the degrees of freedom in modeling time trends were compared in GAM-strict and in GLM-natural splines (Commentary Table 3). (The last comparison was available for only a few of the reports because most investigators limited their analyses to the EPA guidelines for conducting the revised analyses.) The changes were categorized arbitrarily into those judged to be "substantial" ($\geq$ 40% change), to show "some" change (> 10% to < 40%), and to show "little or no" change ($\leq$ 10%).

More detailed information on results is presented in Appendix A, which presents each short report in alphabetical order by surname of the first author. This appendix highlights and compares the original results (GAM-default) with GAM-strict, with GLM-natural splines, with penalized splines and with other methods reported by the investigators. Estimates of the effect of PM concentrations on total mortality from nonaccidental causes, mortality from and hospital admissions for cardiovascular and respiratory diseases, and other selected outcomes are presented.

RESULTS OBTAINED WITH GAM-DEFAULT, GAM-STRICT, AND GLM-NATURAL SPLINES

Overall, tightening convergence criteria and maximum number of iterations in GAM (GAM-strict) decreased GAM-default effect estimates by more than 40% in only two of the studies (*Goldberg and Burnett*[†]*; Ito*). Dominici and colleagues (NMMAPS in this Special Report) also reported substantial changes in the effect estimates from the revised analyses of the 90-cities NMMAPS data. *Goldberg and Burnett* found substantial decreases in effect estimates for total mortality from nonaccidental causes among all Montreal residents. *Ito* found substantial decreases in effect estimates for specific outcomes among Detroit residents. In the remaining studies, a moderate to slight decrease or, rarely, even a slight increase in effect estimates was observed with GAM-strict (*Atkinson et al; Burnett and Goldberg; Fairley; Gold et al; Hoek; Ito; Katsouyanni et al; Klemm and Mason; Le Tertre et al; Mar et al; Moolgavkar; Ostro et al; Samoli et al; Schwartz a,b; Sheppard; Stölzel et al; Zanobetti and Schwartz a,b*) (Commentary Table 1).

For at least one of the outcomes being evaluated, analyzing data with GLM-natural splines was associated with a more pronounced decrease in effect estimates compared with GAM-strict estimates (*Burnett and Goldberg; Goldberg and Burnett; Ito; Katsouyanni et al; Klemm and Mason; Samoli et al; Schwartz a,b; Sheppard; Zanobetti and Schwartz a,b).* *Goldberg and Burnett* reported more pronounced decreases in estimates obtained with GLM, but they modeled time and weather-related variables differently in GAM and GLM, so strict comparability is not possible. Other investigators reported only small decreases in estimates or slight increases for some or most of the outcomes evaluated in each study (*Atkinson et al; Fairley; Gold et al; Hoek; Ito; Le Tertre et al; Mar et al; Moolgavkar; Ostro et al; Schwartz a,b; Stölzel et al; Zanobetti and Schwartz a,b*) (Commentary Table 2). Dominici and Schwartz and their colleagues reported moderate changes in effect estimates obtained with GLM compared with GAM-strict among residents of all cities included in NMMAPS (Dominici et al and Schwartz et al NMMAPS in this Special Report).

Changes in the overall estimates for mortality in the cities included in the Air Pollution and Heath: A European Approach 2 (APHEA2) project were comparable, though somewhat smaller, than those observed in the revised analyses of mortality in the 90 cities in NMMAPS (Dominici et al NMMAPS in this Special Report; *Katsouyanni et al; Samoli*

[†] The short communication reports presented in this Special Report are cited in italics throughout this Commentary.

et al; Schwartz et al NMMAPS in this Special Report; *Zanobetti and Schwartz b*). Overall morbidity estimates for cities that are part of APHEA2 changed little from the original when applying GAM with more stringent conversion criteria, GLM-natural splines, and penalized spline models (*Atkinson et al; Le Tertre et al*) (Commentary Tables 1 and 2). In APHEA2, parametric models yielded overall mortality estimates that were over 30% lower for ambient PM less than 10 μm in aerodynamic diameter (PM_{10}) and black smoke when compared with nonparametric models. When specific estimates were calculated for each of the 29 cities, however, the values both increased and decreased compared with original estimates. When the magnitude of change was compared to the degrees of freedom used to fit the temporal trends, the degrees of freedom selected appeared unrelated to differences observed among cities. For ambient black smoke, for example, the difference in effect estimates (β coefficients) obtained with GAM-default and GLM-natural splines were −1.0% to −122.6% for cities for which time was modeled with approximately 3.5 *df*/yr. For those modeled with 2 *df*/yr, the estimates were −42.3% to +332% (*Katsouyanni et al*). Similar differences were observed for effect estimates of PM_{10}; differences in β coefficients for cities modeled with 3 *df*/yr were −45.7% to +57.3%. The APHEA investigators did not compare the effect of varying the degrees of freedom in city-specific models on estimates of effect obtained with different analytic methods.

RESULTS RELATIVE TO CONTROL FOR TIME

Varying the degrees of freedom in modeling time was evaluated in only 5 short communication reports (*Burnett and Goldberg; Ito; Klemm and Mason; Moolgavkar; Ostro et al*) (Commentary Table 3). In their multicity analysis, *Burnett and Goldberg* found substantial decreases in effect estimates with increasing degrees of freedom for time in the model. They also evaluated the degree of freedom required to model time to drive the model residuals to white noise. They observed that fewer degrees of freedom were required when every sixth day mortality and PM data were included than when all days of observation were included.

In their revised analysis of the replicated Six Cities Study, *Klemm and Mason* found that modeling time with approximately 100 *df* (12 knots/year) in the GLM was associated with the greatest decrease in estimates of effect for all pollutants evaluated. The change in mortality from ischemic heart disease related to the concentration of PM less than 2.5 μm in aerodynamic diameter ($PM_{2.5}$) was 1.8% (95% confidence interval [CI] = 1.1, 2.5; *t* = 5.18) with GAM-strict and 37 *df*; and it was 0.8 (95% CI = 0.0, 1.6; *t* = 1.96) with GLM and more than 100 *df*. Using more than 80 *df* (~10 *df*/yr) with GAM with locally weighted smoothers (LOESS) was

associated with negative estimates of $PM_{2.5}$ effect on total mortality in 2 of the 6 cities evaluated. Using over 200 *df* (~25–40 *df*/yr) with GLM-natural splines resulted in negative effect estimates for 4 of the 6 cities evaluated.

For the most part, increasing degrees of freedom for time decreased the estimates of effect (*Burnett and Goldberg; Ito; Klemm and Mason; Moolgavkar).* This finding was not universal, however, because estimates for some outcomes changed little or increased when degrees of freedom were increased (*Moolgavkar; Ostro et al*).

Moolgavkar reported results from increasing degrees of freedom for time to 100 in the reanalysis of two of the three counties that had been included in his earlier study. Effect estimates changed substantially, moderately, or little depending on the county, outcome, and modeling of variables. Estimates of the effect of PM_{10} and $PM_{2.5}$ at several lags dropped substantially for nonaccidental mortality in Los Angeles County when degrees of freedom for time were increased from 30 to 100 (from ~3.75 *df*/yr to ~12.5 *df*/yr). (For $PM_{2.5}$ at lag 1, the change in mortality was 0.59% [*t* = 1.96] with GAM strict using 30 *df* and 0.10% [*t* = 0.35] using 100 *df* in GAM and −0.01% with GLM using 100 *df*.) For the Cook County data, however, increasing degrees of freedom yielded moderately different values at lags 0 and 1 for PM_{10} and mortality. For example, in Cook County, the observed change in total mortality at lag 1 was 0.43% (*t* = 3.86) per 10 μg/m³ PM_{10} concentration for a model with 30 *df*, 0.38% (*t* = 3.65) for one with 100 *df* in GAM-strict, and 0.37% (*t* = 3.28) for GLM with 100 *df*. Change was moderate in both counties for PM_{10} and $PM_{2.5}$ at lags 0 and 1 and hospital admissions for cardiovascular disease among the elderly. For example, for PM_{10} at lag 0 and admissions related to cardiovascular diseases, the effect estimates for models with 100 *df* were 0.76% (*t* = 6.49) for GAM-default, 0.81% (*t* = 6.89) for GAM-strict, and 0.85% (*t* = 6.76) for GLM-natural splines. Results for lag 1 were similar. (The percent change per 25 μg/m³ increase in PM_{10} concentration for lag 0 from GAM with default settings was 2.1 [*t* = 6.8] for these diagnoses [Moolgavkar 2000].) Overall, changes in effect estimates for lags 2 to 5 were more pronounced and substantial, including for estimates of effect of gaseous pollutants (*Moolgavkar*).

Ostro et al found that increasing degrees of freedom for day and temperature in GAM and GLM affected the effect estimate only slightly for coarse particles and mortality from cardiovascular diseases relative to results obtained with models with fewer degrees of freedom for time and temperature (β = 0.84 [SE = 0.45] with GLM with time modeled with 10 *df*, β = 0.88 [SE = 0.47] with 60 *df* and β = 0.93 [SE = 0.48] with 60 *df* for time and 20 *df* for temperature; GAM-strict corresponding coefficients were 0.90 [SE = 0.44], 0.92 [SE = 0.44] and 0.97 [SE = 0.45]).

Commentary Table 1. Estimates of Effect for GAM-Strict Compared with GAM-Default Estimates[a]

First Author Original/Revised	Health Outcomes	PM Measurements	No. Cities, Location	Degrees of Freedom
Substantial Change (≥ 40%)[b]				
Goldberg	Total mortality	COH, predicted $PM_{2.5}$	Montreal	Time 73.5 (~7 df/yr), temp & pressure 9.2 df
Lippmann/Ito	Total mortality, 1985–1990, 1992–1994; Circulatory mortality, 1985–1990; COPD hosp among elderly	PM_{10}, $PM_{2.5}$; PM_{10}; $PM_{2.5}$	Detroit	Time 12 df/yr, except for resp mortality (1985–1990) 4 df/yr; weather variables ~2.5 df
Some Change (> 10% to < 40%)[b]				
Burnett	Total mortality	PM_{10}, $PM_{2.5}$, $PM_{10-2.5}$	8, Canada	Time 90 day span, temp & pressure 0.5 span, 4018 days, common weather model for all cities
Goldberg	Total mortality among elderly with underlying resp or coronary diseases or congestive heart failure	COH, predicted $PM_{2.5}$	Montreal	Time 73.5 (~7 df/yr), temp & pressure 9.2 df
Lippmann/Ito	Circulatory mortality, 1992–1994; Resp, pneumonia, & heart failure hosp among elderly	PM_{10}, $PM_{10-2.5}$, $PM_{2.5}$	Detroit	Time 12 df/yr, except for resp mortality (1985–1990) 4 df/yr; weather variables ~2.5 df
Klemm	Total mortality, all ages & among elderly; COPD mortality among elderly[c]	PM_{10}, PM_{15}, $PM_{10-2.5}$, $PM_{2.5}$,	6, US	GAM-default 36.8 df (~4.6 df/yr), GAM-strict 51 df (~6.4 df/yr)
Moolgavkar[d]	Total mortality	PM_{10}, $PM_{2.5}$	LA County, CA	Time 100 df (~12.5 df/yr), temp 6 df
Sheppard	Asthma hosp	PM_{10}, $PM_{10-2.5}$	Seattle	Time 64 df (~9 df/yr)
Little or No Change (≤ 10%)[b]				
Atkinson	Asthma & resp hosp	PM_{10}, PM_{13}, BS	8, Europe	City-specific models, same as in original APHEA2 analysis
Fairley	Total, CVD, & resp mortality	PM_{10}, $PM_{10-2.5}$, $PM_{2.5}$	1 county, CA	Time 7 df (~0.9 df/yr), season 12 df (~2.4 df/yr), min temp 3 df, max temp 2 df
Gold	Heart rate variability (ECG tracings)	$PM_{2.5}$	Boston	Temp 2 df
Hoek	Total & cause-specific mortality	PM_{10}, BS	The Netherlands	Time: total 94 df (10.4 df/yr), resp & CVD 60 df (6.67 df/yr)
Katsouyanni	Total mortality	PM_{10}, BS	29, Europe	City-specific models: time ~1.6–4.4 df/yr, temp 2–9 df

Study	Outcome	PM	Location	Model
Lippmann/Ito	Resp mortality, 1985–1990, 1992–1994 Stroke & dysrhythmias hosp among elderly	PM_{10}, $PM_{10-2.5}$, $PM_{2.5}$	Detroit	Time 12 df/yr, except for resp mortality 4 df/yr; weather variables ~2.5 df
Klemm	IHD & pneumonia mortality	$PM_{2.5}$	6, US	GAM-default 36.8 df (~4.6 df/yr); GAM-strict 51 df (~6.4 df/yr)
Le Tertre	Cardiac diagnoses & stroke hosp	PM_{10}, PM_{13}	8, Europe	City-specific models, same as in original APHEA2 analysis
Mar	CVD mortality, age 65–100 years	PM_{10}, $PM_{2.5}$, $PM_{10-2.5}$	1 county, AZ	Time 10 df (~3.3 df/yr), temp 2 df, humidity 2 df
Moolgavkar[d]	Total mortality	PM_{10}, $PM_{2.5}$	LA County, CA	Time 30 df (~3.7 df/yr), temp 6 df
	CVD mortality	PM_{10}, $PM_{2.5}$	LA County, CA	Time 30 df or 100 df
	CVD hosp	PM_{10},	Cook County, IL	Time 100 df
Ostro	CVD mortality	PM_{10}, $PM_{2.5}$, $PM_{10-2.5}$	Valley in CA	Time 10 df (1 df/yr) (after testing 10–60 df)
Samoli	Total mortality	BS	8, Europe	City-specific models: time 12–57 df (3–5 df/yr); temp 2–9 df, humidity 5 df or linear
Schwartz	Total mortality among elderly	$PM_{2.5}$	6, US	Time 36 df (~4 df/yr), temp & humidity ~3.6 df
Sheppard	Asthma hosp	$PM_{2.5}$	Seattle	Time 64 df (~9 df/yr)
Wichmann/Stölzel	Total mortality	$PM_{2.5}$, PM ultrafine, particle number concentration, particle mass concentration	Erfurt, Germany	Time ~4 df/yr
Zanobetti	Total mortality	PM_{10}	10, Europe	City-specific models. Time 3–23 df (1–4.6 df/yr), temp lag 0 2–6 df, temp mean 2–3 days 2–6 df, humidity lag 0 or mean lag 1–2 1.2–2 df

Hosp = hospital admissions; max = maximum; min = minimum; pressure = barometric pressure; resp = respiratory disease; temp = temperature.

[a] Outcome of interest, PM fractions, geographic location, and modeling of time or weather variables also listed.

[b] Changes may vary in direction, especially within the Little or No Change category, where changes may be negative, positive, or both depending on exposure, exposure lags, and outcome.

[c] Klemm: Estimates for mortality from COPD from exposure to $PM_{2.5}$ were higher with GAM-strict.

[d] Moolgavkar: Changes in effect estimates obtained with different models varied according to exposure lag for PM_{10} and $PM_{2.5}$. This table notes results for lags 0 and 1. Estimates changed in a more pronounced manner for lag 2 exposures and were close to zero for most other lags when modeled with 100 df for time.

Commentary Table 2. Estimates of Effect for GLM with Natural Cubic Splines Compared with GAM-Strict Estimates[a]

First Author Original/Revised	Health Outcomes	PM Measurements	No. Cities, Location	Degrees of Freedom
Substantial Change (≥ 40%)[b]				
Burnett	Total mortality	$PM_{2.5}$	8, Canada	Time 90 day span, temp & pressure 0.5 span, 4018 days, common weather model for all cities. GLM 6 knots/yr, temp & humidity 2 df; common & city-specific models vary
Goldberg	Total mortality & mortality among elderly with underlying CVD & resp	COH, predicted $PM_{2.5}$	Montreal	GAM Time 73.5 df (~7 df/yr), temp & pressure 9.2 df. GLM 88–27 df (~8–3 df/yr), temp 0–7 df, humidity 0–7 df, pressure 0–6 df
Lippmann/Ito	Circulatory mortality, 1985–1990	PM_{10}	Detroit	Time 12 df/yr, except resp mortality (1985–1990) 4 df/yr; weather variables ~2.5 df. GLM approximated GAM, weather variables 2 df each
	COPD & stroke hosp among elderly	$PM_{2.5}$		
Klemm	Total mortality, all ages	PM_{10}, PM_{15}, $PM_{10-2.5}$	6, US	GAM strict 30 df (~6.4 df/yr), GLM 38 df (~4.75 df/yr)
	Pneumonia mortality among elderly	$PM_{2.5}$		
Samoli	Total mortality	BS	8, Europe	City-specific models. Time 12–57 df (3–5 df/yr), temp 2–9 df; humidity 5 df or linear
Some Change (> 10% to < 40%)[b]				
Burnett	Total mortality	PM_{10}, $PM_{10-2.5}$	8, Canada	Time 90 day span, temp & pressure 0.5 span, common weather model for all cities. GLM: 6 knots/yr selected for 4018 days, temp & humidity 2 df; common & city-specific models vary. Models with ~670 days of pollutant had fewer knots
Hoek	Total & CVD mortality[c]	PM_{10}	The Netherlands	Time: total mortality 94 df (10.4 df/yr), resp & CVD 60 df (6.67 df/yr)
Lippmann/Ito	Total mortality, 1985–1990, 1992–1994	PM_{10}, $PM_{10-2.5}$	Detroit	Time 12 df/yr, except resp mortality (1985–1990) 4 df/yr; weather variables ~2.5 df; GLM approximated GAM, weather variables 2 df each
	Resp mortality, 1985–1990	$PM_{2.5}$, PM_{10}		
	COPD, stroke & some CVD hosp in elderly	PM_{10}, $PM_{10-2.5}$		
Katsouyanni	Total mortality	PM_{10}, BS	29, Europe	City-specific models: Time ~1.6–4.4 df/yr, temp 2–9 df
Klemm	Total mortality, all ages	$PM_{2.5}$	6, US	GAM-strict ~6.4 df/yr; GLM ~4.75 df/yr
	Total COPD & IHD mortality in elderly			
Le Tertre[d]	Cardiac hosp among elderly	PM_{10}, PM_{13}	8, Europe	City-specific models. Same as in original APHEA2 analysis
Moolgavkar[e]	COPD mortality	PM_{10}	Cook County, IL	Time 30 df (~3.7 df/yr) & 100 df (~12.5 df/yr), temp 6 df
	Total mortality	PM_{10}, $PM_{2.5}$	LA County, CA	
	CVD mortality, COPD hosp	$PM_{2.5}$	LA County, CA	
	COPD hosp	PM_{10}	LA County, CA	
Ostro	CVD mortality	PM_{10}	Valley in CA	Time 10 df (1 df/yr) (after testing 10–60 df)
Sheppard	Asthma hosp	PM_{10}, $PM_{2.5}$	Seattle	GAM & GLM time 64 df (~9 df/yr), GLM temp 4 df

Little or No Change ($\leq$ 10%)[b]

Atkinson	Asthma & resp hosp	PM$_{10}$, PM$_{13}$, BS	8, Europe	City-specific models. Same as in original APHEA2 analysis
Fairley	Total, CVD, & resp mortality	PM$_{2.5}$, PM$_{10}$, PM$_{10-2.5}$	1 county, CA	GAM: time 7 df (~0.9 df/yr), season 12 df (~2.4 df/yr), min temp 3 df, max temp 2 df; GLM: time 7 df (~0.9 df/yr), season 4 df (1.4 df/yr), min temp 3 df, max temp 3 df
Gold	Heart rate variability (ECG tracings)	PM$_{2.5}$	Boston	GAM temp 2 df, GLM temp 3 df
Hoek	COPD & pneumonia mortality	PM$_{10}$, BS	The Netherlands	Time: total mortality 94 df (10.4 df/yr), resp & CVD 60 df (6.67 df/yr)
Lippmann/Ito	Circulatory mortality, 1992–1994; Hosp for some CVD diagnoses among elderly	PM$_{2.5}$, PM$_{10}$, PM$_{10-2.5}$	Detroit	Time 12 df/yr, except resp mortality (1985–1990) 4 df/yr; weather variables span 0.5. GLM approximated GAM, weather variables 2 df each
Le Terrte	Cardiac hosp	PM$_{10}$, PM$_{13}$	8, Europe	City-specific models. Same as in original APHEA2 analysis
Mar	CVD mortality, age 65–100 years	PM$_{2.5}$, PM$_{10}$, PM$_{10-2.5}$	1 county, AZ	GAM time 10 df (~3.3 df/yr), temp 2 df, humidity 2 df; GLM same, except humidity 3 df
Moolgavkar[e]	Total & CVD mortality CVD hosp CVD hosp COPD hosp	PM$_{10}$ PM$_{10}$ PM$_{10}$ PM$_{2.5}$	Cook County, IL Cook County, IL LA County, CA LA County, CA	Time 100 df (~12.5 df/yr), temp 6 df
Ostro	CVD mortality	PM$_{10-2.5}$, PM$_{2.5}$	Valley in CA	Time 10 df (1 df/yr)
Schwartz	Total mortality	PM$_{2.5}$	6, US	GAM: time 36 df (~4 df/yr), temp & humidity 3.6 df; GLM: same as GAM but with 1 extra df added to each variable
Sheppard	Asthma hosp	PM$_{10-2.5}$	Seattle	GAM & GLM: time 64 df (~9 df/yr); GLM temp 4 df
Wichmann/Stölzel	Total mortality	PM$_{10}$, PM$_{2.5}$, PM ultrafine, particle number concentration, particle mass concentration	Erfurt, Germany	Time ~4 df/yr

Hosp = hospital admissions; max = maximum; min = minimum; pressure = barometric pressure; resp = respiratory disease; temp = temperature.

[a] Outcome of interest, PM fractions, geographic location, and modeling of time or weather variables are also listed.

[b] Changes may vary in direction, especially within the Little or No Change category, where changes may be negative, positive, or both depending on exposure, exposure lags, and outcome.

[c] Hoek: Estimates for exposure to PM$_{10}$, lagged 0–6 days, obtained with GLM-natural splines were smaller than those obtained with GAM-strict for total mortality and higher for CVD mortality.

[d] Le Terrte: Estimates obtained with GLM-natural splines were larger than those obtained with GAM-strict.

[e] Moolgavkar: Changes in effect estimates obtained with different models varied according to exposure lag for PM$_{10}$ and PM$_{2.5}$. This table notes results for lags 0 and 1. Estimates changed in a more pronounced manner for lag 2 exposures and were close to zero for most other lags when modeled with 100 df for time.

Commentary Table 3. Estimates of Effect After Increasing Degrees of Freedom in Modeling Time[a]

First Author Original/Revised	Health Outcomes	PM Measurements	No. Cities, location	Degrees of Freedom
Substantial Change (≥ 40%)[b]				
Burnett	Total mortality	$PM_{10-2.5}$	8, Canada	GLM time 1–12 knots/yr
Lippmann/Ito	Total mortality, 1992–1994 Pneumonia hosp, 1992–1994	PM_{10} PM_{10}	Detroit	GLM: varying lags, weather variables and increasing *df* for time, temp & humidity
Klemm	Total mortality COPD, pneumonia & IHD mortality in elderly[d]	$PM_{2.5}$, PM_{10}, PM_{15}, $PM_{10-2.5}$ $PM_{2.5}$	6, US	GLM time < 4 to 12 knots/yr
Moolgavkar[c]	Total mortality[d] COPD mortality[d]	PM_{10}, $PM_{2.5}$ $PM_{2.5}$	LA County, CA	GAM & GLM time 30 *df* (~3.7 *df*/yr) & 100 *df* (~12.5 *df*/yr), temp 6 *df* GAM 100 *df* & GLM 100 *df* only
Some Change (> 10% to < 40%)[b]				
Burnett	Total mortality	$PM_{2.5}$	8, Canada	GLM time 1–12 knots/yr
Moolgavkar[c]	Total & CVD[d] mortality CVD mortality CVD & COPD hosp	PM_{10} PM_{10}, $PM_{2.5}$ PM_{10}, $PM_{2.5}$	Cook County, IL LA County, CA LA County, CA	GAM only, time 30 *df* (~3.7 *df*/yr) & 100 *df* (~12.5 *df*/yr), temp 6 *df*
Little or No Change (≤ 10%)[b]				
Moolgavkar[c]	COPD mortality CVD hosp	PM_{10} PM_{10}	Cook County, IL	GAM time 30 *df* (~ 3.7 *df*/yr) & 100 *df* (~12.5 *df*/yr), temp 6 *df*; GLM 100 *df* only
Ostro	CVD mortality	$PM_{10-2.5}$	Valley in CA	Time 10–60 *df* (1–6 *df*/yr), temp 20 *df*

Hosp = hospital admissions; temp = temperature.

[a] Outcome of interest, PM fractions, geographic location, and modeling of time or weather variables are also listed.

[b] Changes may vary in direction, especially within the Little or No Change category, where changes may be negative, positive, or both depending on exposure, exposure lags, and outcome.

[c] Moolgavkar: Changes in effect estimates obtained with different models varied according to exposure lag for PM_{10} and $PM_{2.5}$. This table notes results for lags 0 and 1. Estimates changed in a more pronounced manner for lag 2 exposures and were close to zero for most other lags when modeled with 100 *df* for time.

[d] Moolgavkar: Estimates increased and/or decreased substantially (≥ 40%) or moderately (39%–11%) when modeling time with 100 *df* compared with 30 *df*.

OTHER FACTORS

Ito in Detroit and *Goldberg and Burnett* in Montreal explored weather factors in more depth than did other investigators. These investigators found that estimates of the effect of air pollutants on total mortality were sensitive to modeling of weather terms and temporal trends.

Ito evaluated the association of PM_{10} with mortality and hospital admissions for pneumonia among the elderly by applying models with varying degrees of freedom for time (1/yr to 24/yr) and by varying lags (lag 0, average of lags 1–3) and degrees of freedom (2–6 *df*) for temperature and dew point (*Ito*, Figures 9 and 10). For pneumonia, for all models that controlled for weather variables, the lowest estimate was obtained when the models included 4 *df*/yr for time (β ~0.0006–0.0017); slightly higher levels were obtained with 12 *df*/yr. The highest estimate was obtained with the model that included 24 *df*/yr for time and did not adjust for any weather factors (β ~0.0045) (*Ito*, Figure 10). *Ito* found decreases in coefficients for total mortality and pneumonia when weather variables averaged and lagged over different days were included in the models (*Ito*, Figures 9 and 10). During this period of study, 1992 to 1994, PM_{10} concentration was available for only 490 days, a limitation possibly affecting the dependence of estimates of effect on number of degrees of freedom used to model time. Data on weather variables were available for 1096 days (Lippmann et al 2000).

Goldberg and Burnett found a pronounced decrease in PM effect estimates when applying GLM-natural splines with a parameter smooth function for temperature. The estimates were highly sensitive to altering the way temperature was specified in the model (Table 8). The authors did not observe substantial decreases in the effect estimates for coefficient of haze and predicted $PM_{2.5}$ on total nonaccidental mortality with the use of GLM when they included indicator variables for higher temperatures, instead of a smooth function in the model (Tables 9 and 10).

In addition, in their multicity analysis on the effect of air pollution on total mortality, *Burnett and Goldberg* conducted tests for heterogeneity among 8 Canadian cities. Tests indicated presence of heterogeneity with LOESS models, but indicated no heterogeneity with natural spline models.

Dominici et al, Schwartz b, and *Zanobetti and Schwartz b* conducted revised analyses to evaluate mortality displacement. *Dominici et al* evaluated the effect on total mortality of total suspended particles during the years between 1974 and 1988 in Philadelphia. Although effect estimates were halved when analyzing data with GAM-strict and GLM-natural splines, relative risks were nevertheless higher with longer time scales than with shorter time scales. *Schwartz b*, in his revised analysis of ambient $PM_{2.5}$ and mortality displacement in Boston, found the same patterns

as reported previously (Schwartz 2000). Effect estimates for mortality from ischemic heart disease and pneumonia increased with longer time scales, while effect estimates for chronic obstructive pulmonary disease decreased with time. *Zanobetti and Schwartz b* reevaluated results on mortality displacement from 10 European cities. Effect estimates from fourth degree polynomial distributed lag models for ambient PM_{10} (although lower when obtained with GAM-strict and with penalized splines than with GAM-default) were higher than comparable estimates obtained from mean 2-day lag models.

OVERALL RESULTS

Differences in effect estimates varied substantially across and within studies. Overall, the revised analyses using GAM with more stringent convergence criteria and iterations and GLM-natural splines resulted in lower estimates, but largely continued to find an association of PM with mortality (*Burnett and Goldberg; Dominici et al; Katsouyanni et al; Samoli et al; Schwartz b; Zanobetti and Schwartz b*) and morbidity, in particular for cardiovascular and respiratory diseases (*Atkinson et al; Fairley; Gold et al; Hoek; Ito; Le Tertre et al; Ostro et al; Schwartz a; Sheppard; Zanobetti and Schwartz b*). As in earlier analyses, the effect estimate was larger among those who were 65 years of age and older (*Fairley; Gold et al; Goldberg and Burnett; Ito; Le Tertre et al; Mar et al; Moolgavkar; Schwartz a*). The impact of various sensitivity analyses, when these were performed, differed across the studies. No significant impacts were seen in some (*Ostro et al*), whereas in others, alternative modeling of time (*Klemm and Mason; Moolgavkar*) and weather factors (*Goldberg and Burnett; Ito*) resulted in substantial changes.

METHODS APPLIED IN REVISED ANALYSES

STRICTER CONVERGENCE CRITERIA

Since 1996, several published papers on the effects of daily fluctuations in air pollution on daily mortality counts have reported results obtained with GAMs. More specifically, Poisson regression methods have been used in which air pollution and other risk factors and/or confounders are fitted linearly, along with smooth functions of time or weather. The effect of calendar time is likely to be nonlinear because season and time-dependent unmeasured risk factors (such as viral epidemics) will not be linear. The effect of some weather variables (eg, temperature) has also been found to be markedly nonlinear in numerous studies. These possibly nonlinear effects have

been modeled with nonparametric smooth functions (eg, LOESS, smoothing splines) as part of a semiparametric Poisson regression model.

Most of the published papers on this subject, and all of those included in this report, have used the S-Plus statistical package. The S-Plus implementation of this modeling procedure involves two iteration loops: One is the global iteration (the outer loop) on the overall iteratively reweighted regression procedure and the other is an inner Gauss-Seidel type iterative procedure (the backfitting algorithm) that cycles through the estimation of smooth functions (the inner loop). Until recently, the GAM function in S-Plus used 10^{-3} as the default convergence criterion for both the outer and inner loops along with 10 as a maximum number of iterations. Note that when there is only one nonparametric term in the Poisson model, the backfitting algorithm is not needed because a closed form solution is available for this special case (see Hastie and Tibshirani 1990 for details). Thus only the outer loop of the procedure ought to be relevant. But S-Plus uses the backfitting algorithm even for this special case. Hence, depending on the degree of collinearity between the parametric and nonparametric terms in the model, the adequacy of the convergence criterion for the inner backfitting loop may be of concern.

In May 2002, Dominici and colleagues (2002) reported that these default convergence criteria might not be strict enough to allow a fully complete estimation process. They suggested stricter convergence criteria for both the inner and outer loops, leading to a new model that we have termed *GAM-strict*. In the NMMAPS 90 cities mortality study, the use of GAM-strict reduced the combined effect estimate across cities by approximately 40% (Dominici et al 2002), compared to GAM-default estimates.

The short communication reports included in this HEI Special Report comprise results obtained with the stricter convergence criteria. The effect of using GAM-strict on these revised analyses does not appear to be consistent. While some of the studies appear to be relatively more affected by the new convergence criteria, others appear to be insensitive to the change in the convergence criteria.

Differences in effects of GAM-strict on the various studies reflect how close GAM-default took the data to convergence. This depends on factors related to the data and to the possibly subtle differences in the specification and selection of the final models. It is not possible to be definitive as to how various factors affect convergence, but the following have been proposed as relevant:

1. *Number of nonparametric terms*: When only calendar time is fitted nonparametrically, modeling could avoid the backfitting algorithm: within a given iteration of the reweighted least squares procedure, this semiparametric regression problem has a closed form solution. So, one would expect only the studies that use multiple nonparametric terms to be affected by the possible inadequacy of GAM-default. However, the S-Plus implementation of models that have a single nonparametric term still uses the backfitting algorithm instead of the closed form solution (T Hastie, personal communication, January 2003). Thus, even with just one nonparametric term in the model, default S-Plus convergence criteria may be inadequate. Further, when multiple nonparametric terms are included in the model, the problem with GAM-default may be exacerbated (see 2 below).

2. *Degree of smoothing*: Under some circumstances, models with few degrees of freedom in nonparametric terms would be closer to their correct converged form under the inadequate default convergence criteria than would curves with many degrees of freedom. For example, this might occur if S-Plus began its iterative cycle with nonparametric terms set at linear (a common procedure) and they then became progressively but slowly more wiggly with each iteration.

3. *Degree of collinearity*: When the nonparametric term(s) are correlated with each other and/or with the parametric terms in the model, the estimation process and the resulting estimates become unstable. This phenomenon, usually termed *concurvity* in the nonparametric regression literature, may cause the iterative process (outer and/or inner loops) to be unstable, possibly preventing convergence from being reached quickly.

The first two of these factors are determined by modeling decisions by the analysts, but the last is determined by the data (specific to the city and to outcome). Further, measures of collinearity (the last factor) are not published: so evaluating the importance of this factor, or controlling for it when considering the first two factors, is impossible.

In almost all of these revised analyses, time, temperature, and humidity were modeled nonparametrically, so number of nonparametric terms (no. 1 above) can account for little variation in the effect estimates due to use of stricter criteria. Nevertheless, because (unmeasured) variation in collinearity (no. 3 above) could cause differences between cities in the extent of change in results obtained from the new analysis, it is hard to identify whether degree of smoothing (no. 2 above), which does vary between studies, is also important.

The proportional change in the mean estimate from the NMMAPS mortality study was not markedly different in models with more degrees of freedom (Dominici et al NMMAPS in this Special Report, Figure 10; Samet et al 2000, Figure A.1)—although the former shows GLM-natural spline results). Further, *Hoek* found little change between GAM-strict and GAM-default despite using relatively little smoothing (9 *df*/yr). *Goldberg and Burnett* found quite a large change with relatively smooth nonparametric functions, but they modeled variables differently than in the GAM analyses. Thus, degree of smoothing may be one determinant of change on making convergence criteria stricter, but, if so, it is not a dominant one.

ACCURACY OF STANDARD ERRORS

The GAM software uses a standard error estimator that is only valid when no correlation exists between the air pollution term and any of the regression variables. In the presence of correlation, the GAM standard error of the pollution effect underestimates the true standard error. This approximation is clearly documented. Ramsay and colleagues (2003) noted that it is particularly poor when the effect estimate is small and when the time series of PM measurements correlates substantially with several smooth functions being fitted (the latter is often called concurvity). Both of these problems plague the usual models for fitting time series of mortality and morbidity and their dependence on PM_{10}, ozone, and other pollutants.

Research is ongoing to find a correct method for estimating the standard error in GAM models with correlation among the confounders. (Francesca Dominici, Aidan McDermott, and Trevor Hastie recently presented a technical report describing a new program [gam.exact] to estimate accurate standard errors using nonparametric smoothing functions in GAM in S-Plus: "Issues in semiparametric regression with applications in time series models for air pollution and mortality" [*http://biosun01.biostat.jhsph.edu/~fdominic/*].) In the meantime, GLM with natural splines or penalized cubic splines as implemented in R do provide (asymptotically) correct estimates of the standard error of the estimated pollution effect. In the R statistical package, GAMs are fit using GLM with penalized splines, and the amount of smoothing is chosen by generalized cross-validation. This method also gives asymptotically correct standard errors.

The reanalysis of mortality in NMMAPS II by Dominici and colleagues (in this Special Report) uses GLM-natural splines estimates, and the morbidity analysis by Schwartz and colleagues (in this Special Report) also computes estimates with penalized splines. Most of the revised analyses computed GLM-natural splines and the overwhelming

pattern is that the standard errors under this model are somewhat larger than the standard errors under GAMs regardless of whether the GAM uses strict convergence. The convergence setting does not seem to affect the GAM estimates of standard error. The increase in standard error is modest in nearly all cases and has a much smaller effect on the overall conclusions of each study than the change in pollution effect estimate. Thus, while it is important that a correct standard error estimate be implemented, the most dramatic changes in conclusions here are due to the changing estimate of pollution effect.

All the estimates of standard error discussed in the preceding two paragraphs assume the usual overdispersed Poisson model. When the dependence of the error variance on the covariates is not simply through the mean function, none of the preceding software implementations provides correct standard errors; rather, software that computes a so-called robust variance is required (Wood 2003).

One important consequence of the increase in estimated standard errors occurs in multicity studies, such as NMMAPS and APHEA, where the goal is to estimate a national or European effect of pollution on mortality or morbidity. The city-specific standard errors in these studies are larger than they were in the initial analyses, with the result that the variation among cities is less pronounced; thus there is more statistical evidence for combining city-specific estimates using a random effects model without effect modifiers for each city. This increase in the standard errors in turn highlights the difficulty of interpreting the "national" or "Eastern or Western European" average.

SENSITIVITY OF FINDINGS TO ALTERNATIVE MODELS

A potentially useful consequence of the discovery that GAM estimates and standard errors might be in error has been that a number of other smoothing methods have been used to reanalyze the same data. These new analyses allow some investigation of the sensitivity of results to the use of alternative models. Most investigators performed their revised analyses using GLM-natural splines. Several authors also used penalized splines and case-crossover matching. Like GAM, each of these methods attempts to prevent confounding of the relation between air pollution and the response of interest, say mortality, by controlling for the nonlinear effects of time, temperature, and perhaps, humidity. The methods all assume that the effect of these covariates can be summarized by a smooth dose-response function, which must be estimated from the data. Some issues are specific to modeling the effects of time and the use of the case-crossover matching for this purpose. These are discussed in some depth in the HEI Commentary to the

NMMAPS II revised analyses elsewhere in this Special Report. Because case-crossover matching is not fundamentally different from the other three smoothing methods, we do not consider this method further. Each of the other methods is associated with a free parameter, the degrees of freedom, which determines how wiggly that function is allowed to be for each smooth function. A few investigators evaluated in greater detail the effect of different amounts of smoothing on their results.

Sensitivity to Smoothing Models

The GLM-natural splines model, the GAM, and the penalized spline model are defined by model-specific sets of smooth functions. When the degrees of freedom are the same, the size of the model-specific sets of functions are the same under each of the three models. However, the model-specific sets of functions differ somewhat between models. Therefore one would not expect the estimated smooth functions obtained under the three different models to agree precisely. One hopes, however, that all of the sets contain a sufficiently rich class of functions that the actual dose-response function for a given variable, say temperature, can be approximated sufficiently well by any of the sets. Thus confounding can be adequately controlled and the estimated effect of air pollution on mortality would be similar under each model. A survey of the results comparing GLM-natural splines, GAM (with stringent convergence criteria), and penalized splines fit with the same degrees of freedom demonstrates that the three methods often provide estimates of the PM relative risk that differ by no more than 30% although substantially larger differences are occasionally seen. Should we consider these effect estimates to be similar?

What is meant by "a similar effect of air pollution" depends critically on the magnitude of the air pollution effect. The increase in daily (overall) mortality due to a 10 $\mu g/m^3$ rise in PM concentration has been estimated to be on the order of 0.2% to 0.8%. The increase in death from cardiac or respiratory related causes can be 4 to 5 times as large. As the total between-day variation in PM over a short time interval is usually no more than 50 $\mu g/m^3$, the effect of variation in PM on daily mortality is at most a few percent. On the other hand, effects of confounders such as extremes in temperature may have effects on daily mortality that are several times as large (Curriero et al 2002). Given the small size of the pollution effect, a variability between methods up to 30% seems quite expected, although the few substantially larger differences are somewhat surprising and should be further investigated. In particular the sensitivity of the results to using GLM-natural splines versus GAM to model the temperature smooth in *Goldberg and Burnett*'s reanal-

ysis of Montreal data is especially striking. This difference between their results and those of others, however, is likely to be due to modeling time-related and weather-related variables differently in GAM and GLM-natural splines models.

Sensitivity to Degrees of Freedom

Perhaps a more important issue is the sensitivity of all methods to the degrees of freedom allotted. The concerns are as follows. Suppose that (1) the effect of, say temperature or some other weather variable, is characterized by a smooth function with many wiggles and (2) daily pollution is highly correlated with temperature. Then by restricting the degrees of freedom in our temperature smooth, the nonlinear effects of temperature would be falsely ascribed to the pollution variable even under the null hypothesis of no pollution effect. Thus to decrease the potential for confounding bias, it is tempting to afford many degrees of freedom to each potential confounding variable, but there is a drawback to this solution. When these degrees of freedom are not needed because the true dose response is not very wiggly, the result will be a much less efficient estimate of the pollution effect. This loss of efficiency may compromise our ability to detect a true but small pollution effect.

Further, consider the effect if temperature is given many degrees of freedom and if pollution is measured with error or if an improper pollution history is used as the pollution covariate. For instance, the previous day PM_{10} might be used instead of the biologically relevant measure of, for example, the average of the previous three days. Then the result would be that part of any true pollution effect would be incorrectly ascribed to the temperature variable. Although the measurement error can be corrected, at least in part, with statistical methods that correct for measurement error (Zeger et al 2000; Zeger and Diggle 2001), there remains no altogether satisfactory way to choose the appropriate degrees of freedom: The first need to allow many degrees of freedom to assure no confounding bias under the null hypothesis stands opposed to the second need not to include too many degrees of freedom to preserve the statistical power to detect pollution effects. The empirical data cannot determine the optimal trade-off between these conflicting needs, and it is difficult to use an a priori biological or meteorologic knowledge to determine the optimal trade-off.

Few studies have addressed the impact on the estimated pollution effect caused by the degrees of freedom used to model potentially confounding variables and time (Samet et al 1997). NMMAPS investigators show results for doubling and halving their baseline degrees of freedom for time and weather smooth functions with an increase in the PM_{10} coefficient on reducing degrees of freedom, and a

small decrease in the coefficient and small increase in standard error with increasing degrees of freedom (Samet et al 2000; Dominici et al NMMAPS in this Special Report, Figure 10). A few other authors have explored sensitivity of results to a range of degrees of freedom. *Fairley* investigated the sensitivity of results to the choice of degrees of freedom for a one (trend) or two (trend and season) smooth functions of time. *Moolgavkar* considered the impact of 30 and 100 *df* on the smooth function of time (roughly amounting to 3 *df*/yr and 12 *df*/yr). Only *Ito* and *Klemm and Mason* modeled time more finely than about 1 *df*/month. *Klemm and Mason* found a modest increase in standard error with GLMs, and a decrease in coefficient that rounded to about zero at this level of control for both GAM and GLM. *Ito* found decreases in coefficients only when weather variables that averaged and lagged over different days were included in the models.

Approaches to Choosing Degrees of Freedom

In the air-pollution literature, three methods have been used to determine the optimal degrees of freedom, two purely statistical and one substantive.

Statistical Methods Control of confounding by discrete covariates can be handled by exact stratification on levels of the covariate. In contrast, control of confounding by either continuous covariates or discrete covariates with many ordered levels (eg, temperature measured to the nearest degree Celsius) requires a smoothing method and a method to choose degrees of freedom. Common statistical criteria used to choose the degrees of freedom are to minimize the Akaike information criterion (AIC) (*Atkinson et al; Burnett and Goldberg; Fairley; Goldberg and Burnett; Hoek; Ito; Katsouyanni et al; Le Tertre et al; Mar et al; Ostro et al; Samoli et al; Stölzel et al; Zanobetti and Schwartz a,b)* or to make the empirical autocorrelation function of the residuals equal to zero (*Fairley; Ito; Schwartz a,b*). (Under this theory, too few degrees of freedom results in positive empirical autocorrelation and too many degrees of freedom to negative empirical autocorrelation.)

Ritov and Bickel (1990) have shown, however, that for any continuous variable, no strictly data-based (ie, statistical) method can exist by which to choose a sufficient number of degrees of freedom to insure that the amount of residual confounding due to that variable is small. This means that no matter what statistical method one uses to select the degrees of freedom, it is always logically possible that even if the true effect of pollution is null, the estimated effect is far from null due to confounding bias. Indeed, Ritov and Bickel (1990) proved that no strictly data-based method can guarantee to yield either (1) an estimator of the

pollution effect with small bias or (2) a CI for the pollution effect that is both narrow enough to be useful for risk assessment and yet wide enough to guarantee that it covers the true effect at least 95% of the time. In this sense, estimation of the pollution effect is an ill-posed problem that can only be solved by addition of accurate external a priori information restricting the maximum wiggliness of the regression curve for the effect of the potential continuous confounder on either mortality or pollution.

In Appendix B, we use AIC to illustrate this general conclusion of Ritov and Bickel. Specifically, we provide explicit (extreme) examples that demonstrate that choosing the degrees of freedom by AIC does not logically guarantee adequate control of confounding. Indeed, selection of confounding variables by AIC is closely related to selection of variables by forward selection, in which a new variable is added to a regression model only when it adds to the fit of the current model at some prespecified level of significance. The epidemiologic literature has long recognized that forward selection methods can fail to provide adequate control for confounding (Robins and Greenland 1986). Similar problems arise in the use of methods based on minimizing autocorrelation or tests for departure from white noise in residuals. Note that we are not arguing that the methods used in the air-pollution literature to select the degrees of freedom have in fact selected too few degrees of freedom in any particular study. Rather, we are arguing that no statistical method can guarantee that sufficient degrees of freedom have been selected to provide adequate control for confounding.

Substantive Methods For explicit potential confounding variables, in particular temperature, existing empirical and biological knowledge on the nature of effects can inform the choice of degrees of freedom. Moreover, such knowledge can also inform the choice of lag structure and the addition of interactions. The well-replicated and highly biologically plausible finding of a U or J shaped relation of temperature and mortality suggests using sufficient degrees of freedom to allow for such nonlinearity—a requirement respected by all investigators. Knowledge concerning the lag structure of temperature–mortality relations seems less well respected in the air pollution studies, however. While the consequences of hot temperatures are widely found to be quite prompt—lags of not more than a day or two—those of cold temperatures are delayed. In Europe, the effects of low temperatures on mortality have been shown to persist for at least two weeks and probably longer (Mackenbach et al 1993; Huynen et al 2001; Hennessey et al 2002). Models investigating air pollution effects have only occasionally (Keatinge and Donaldson 2000) included such long-lagged temperature effects.

Although such long-lagged low temperature effects are less evident in North America (Curriero et al 2002), evidence has shown effects lagged up to one week (Ferreira Braga et al 2001). Thus models that include longer temperature lags should be investigated. Indeed, such models should be investigated even if the lagged temperature variables are not statistically significant and/or are not selected by AIC.

Less has been published on the estimated effect of multiday interactions between temperature and mortality. Empirical data during deadly heat waves has shown, however, a profound interaction between very high temperatures and the number of consecutive days the temperature remains high. Further, elementary biology can explain such interactions: during a heat wave, frail individuals do not get the chance to reestablish water, salt, or temperature balance for many days. In the air pollution literature, a number of authors have controlled for potential confounding due to such interactions at high temperature by excluding extreme heat wave periods from the analysis. But perhaps more subtle multiday interactions affect mortality under less extreme conditions? Models that include multiday interactions should be investigated.

Why the emphasis on adequacy of the lag and interaction within the structure of models in this discussion of GAMs and their alternatives? First, like temperature, inadequate lag and interaction structure may leave complex and apparently implausible curves in remaining terms. Thus more work is left for these curves, possibly more than they can deliver even in implausible models. Second and more generally, the flexibility of nonparametric and semiparametric functions to be nonlinear can obscure their limitation in reflecting other complexities (such as lags and interaction of terms).

A final generic problem with relying on biological knowledge to choose the appropriate degrees of freedom is that biological knowledge is often used in too narrow a way: no one has the vision to imagine all possible biological mechanisms. An example is the biological logic that temperature should have a J or U impact (ie, concave upward) rather than a multimodal dose-response curve on mortality. Schwartz and colleagues (NMMAPS in this Special Report) show that as they added more degrees of freedom to a GAM model for temperature, multimodal dose-response functions were estimated. This remains the case even when they put both the current day and the previous day's temperature in the model as separate additive smoothing functions. They argue that fewer degrees of freedom should be used for temperature because the resulting multimodal dose-response curves are biologically implausible.

A goodness-of-fit test might show that the data rule out a concave dose-response function while supporting a multimodal response. (Schwartz et al do not report the results of such a test.) But a multimodal temperature dose-response conflicts with strong biological priors. The results might be consistent with lack of multimodality if the effect of weather (say, temperature and humidity) over the past week to ten days on mortality were accurately modeled. Alternatively, perhaps there is a nonadditive interaction between the temperature and/or the humidity over the past several days or even weeks. Such interactions on the fit of a misspecified GAM model using only the past 2 days of temperature (without an interaction between the days or lagged temperature beyond two days) could well induce an apparent multimodal dose-response. In this instance, if a GAM model with many degrees of freedom were replaced by unimodal models with fewer degrees of freedom, then additional confounding bias would result if pollution levels were also predicted by multiday interactions between temperature and humidity. If the empirical multimodality of the GAM is not explainable by chance, it should be seen as possible evidence that the model was misspecified, possibly because it failed to include multiday interaction terms. In such a case, the merits of a GAM model including such interactions should be investigated. When a model without interactions is used, a multimodal GAM may control confounding better than a unimodal model despite its apparent biological implausibility.

Incompleteness of lag, interaction structure of the model, or other omitted covariates, might also explain the rapidly increasing effects of temperatures above 15°C found by *Goldberg and Burnett* and argued by them to be biologically implausible. To use the argument of implausibility to impose a higher threshold for a heat effect, in the face of evidence from data that such a model fits worse, seems inappropriate to this Panel although it does seem reasonable to explore these findings further.

Our conclusion is that it is better to control for confounding by time, temperature and humidity by first using models with many degrees of freedom, many past days, and many interactions. Later the models can be simplified on the basis of substantive biological arguments (1) if and when necessary to reduce standard errors and then (2) only when the simplifications are not contradicted by other tests.

CONCLUSION TO METHODS

The Panel believes that adding more lagged days and between-day interaction terms to control for the possibly complex nonlinear effects of weather may have a greater influence on the results of time-series analyses than the choice among GAM, GLM-natural splines, and penalized

spline models. The choice of the degrees of freedom for any particular smooth function is likely to have intermediate influence. No criteria for model choice based on data, including for degrees of freedom in smooth terms, can guarantee optimal control of confounding. Careful biological arguments together with more detailed comparisons of climates that differ in the correlation structure between the air-pollution and weather time series are required to more accurately separate the effects of weather from the effects of pollution. Models that include more lagged days and between-day interaction terms will have less power to detect true pollution effects, and thus larger data sets with more accurate pollution measurements may be needed. Data with pollution measured on only one day in six may not allow such stringent confounder control while at the same time avoiding loss of pollution effect due to inability to model effects distributed over lags of several days.

CONCLUSIONS AND RECOMMENDATIONS

Based on its review, the Panel reached the following conclusions:

- In response to the needs of the regulatory and scientific communities, the original investigators in revising their analyses have completed a great deal of work in a short period of time. As was the case with the findings of the original studies, the revised findings will continue to help inform regulatory decisions regarding PM.

- The overall impact of revising these studies include:

  ◦ While the number of studies showing an association of PM with mortality was slightly smaller, the PM association persisted in the majority of studies.

  ◦ In some of the large number of studies in which the PM association persisted, the estimates of PM effect were substantially smaller.

  ◦ In the few studies in which investigators performed further sensitivity analyses, some showed marked sensitivity of the PM effect estimate to the degree of smoothing and/or the specification of weather.

- The impact of using more appropriate convergence criteria on the estimates of PM effect in the revised analyses varied greatly across the studies. In some studies, stricter convergence criteria had little impact, and in a few the impact was substantial. In no study

were conclusions based on the original analyses changed in a meaningful way by the use of stricter criteria. Although several explanations for variability of this effect were considered, including the degree of temporal smoothing used in the original analyses, the number of smoothed terms in the models, and the degree of nonlinear collinearity (concurvity) among the smoothed terms in the models, the relative impact of these and other explanations is unclear.

- In most studies, parametric smoothing approaches used to obtain correct standard errors of the PM effect estimates produced slightly larger standard errors than the GAM. The impact of these larger standard errors on level of statistical significance of the PM effect was minor.

- In general, the original PM effect estimates were more sensitive to the method used to account for temporal effects than to changing the convergence criteria. Further, in the few studies in which this was examined, many estimates of effect were more sensitive to the degree of smoothing of temporal effects than either the use of stricter convergence criteria or the method used to account for temporal effects. While in some studies the original effect estimates were largely insensitive to either the method or degree of smoothing, in several studies the changes were substantial enough to result in meaningful changes in the study conclusions. In those few studies in which qualitative conclusions were changed as a result of the different approaches to smoothing, the revised conclusions indicated no effect of PM.

- While the alternative approaches used to model temporal effects in the revised analyses addressed the problems of obtaining incorrect effect estimates and standard errors when using the GAMs software, at this time none can be recommended as being strongly preferred over another for use in this context.

- Neither the appropriate degree of control for time in these time-series analyses, nor the appropriate specification of the effects of weather, has been determined. This awareness introduces an element of uncertainty into the time-series studies that has not been widely appreciated previously. At this time, in the absence of adequate biological understanding of the time course of PM and weather effects and their interactions, the Panel recommends exploration of the sensitivity of these studies to a wider range of alternative degrees of smoothing and to alternative specifications of weather variables in time-series models.

REFERENCES

Atkinson RW, Anderson HR, Sunyer J, Ayres J, Baccini M, Vonk JM, Boumghar A, Forastiere F, Forsberg B, Touloumi G, Schwartz J, Katsouyanni K. 2003. Acute effects of particulate air pollution on respiratory admissions. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Burnett RT, Goldberg MS. 2003. Size-fractionated particulate mass and daily mortality in eight Canadian cities. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Curriero FC, Heiner KS, Samet JM, Zeger SL, Strug L, Patz JA. 2002. Temperature and mortality in 11 cities in the eastern United States. Am J Epidemiol 155:80–87.

Dominici F, Daniels M, McDermott A, Zeger SL, Samet JM. 2003a. Shape of the exposure–response relation and mortality displacement in the NMMAPS database. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Dominici F, McDermott A, Daniels M, Zeger SL, Samet JM. 2003b. Mortality among residents of 90 cities. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiol 156:193–203.

Fairley D. 2003. Mortality and air pollution for Santa Clara County, California, 1989–1996. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Ferreira Braga AL, Zanobetti A, Schwartz J. 2001. The time course of weather-related deaths. Epidemiology 12:662–667.

Gold DR, Schwartz J, Litonjua A, Verrier R, Zanobetti A. 2003. Ambient pollution and reduced heart rate variability. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Goldberg MS, Burnett RT. 2003. Revised analysis of the Montreal time-series study. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Hastie TJ, Tibshirani RJ. 1990. Generalized Additive Models. Monographs on Statistics and Applied Probability 43, Chapman & Hall, New York NY.

Hennessy E. 2002. Air pollution and short term mortality. BMJ 324:691–692.

Hoek G. 2003. Daily mortality and air pollution in The Netherlands. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Huynen MMTE, Martens P, Schram D, Weijenberg MP, Kunst AE. 2001. The impact of heat waves and cold spells on mortality rates in the Dutch population. Environ Health Perspect 109:463–470.

Ito K. 2003. Associations of particulate matter components with daily mortality and morbidity in Detroit, Michigan. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Katsouyanni K, Touloumi G, Samoli E, Petasakis Y, Analitis A, Le Tertre A, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. 2003. Sensitivity analysis of various models of short-term effects of ambient particles on total mortality in 29 cities in APHEA2. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Keatinge WR, Donaldson GC. 2001. Mortality related to cold and air pollution in London after allowance for effects of associated weather patterns. Environ Res 86:209–216.

Klemm RJ, Mason R. 2003. Replication of reanalysis of Harvard Six-City mortality study. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Le Tertre A, Medina S, Samoli E, Forsberg B, Michelozzi P, Boumghar A, Vonk JM, Bellini A, Atkinson R, Ayres JG, Sunyer J, Schwartz J, Katsouyanni K. 2003. Short-term effects of particulate air pollution on cardiovascular diseases in eight European cities. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Lippmann M, Ito K, Nádas A, Burnett RT. 2000. Associations of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations. Research Report 95. Health Effects Institute, Cambridge MA.

Mackenbach JP, Looman CWN, Kunst AE. 1993. Air pollution, lagged effects of temperature, and mortality: The Netherlands 1979–87. J Epidemiol Community Health 47:121–126.

Mar TF, Norris GA, Larson TV, Wilson WE, Koenig JQ. 2003. Air pollution and cardiovascular mortality in Phoenix, 1995–1997. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Moolgavkar SH. 2000. Air pollution and hospital admissions for diseases of the circulatory system in three US metropolitan areas. J Air Waste Manage Assoc 50:1199–1206.

Moolgavkar SH. 2003. Air pollution and daily deaths and hospital admissions in Los Angeles and Cook counties. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Ostro BD, Broadwin R, Lipsett MJ. 2003. Coarse particles and daily mortality in Coachella Valley, California. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Ramsay TO, Burnett RT, Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology 14:18–23.

Ritov Y, Bickel P. 1990. Achieving information bounds in non- and semi-parametric models. Ann Stat 18:925–938.

Robins JM, Greenland S. 1986. The role of model selection in causal inference from nonexperimental data. Am J Epidemiol 123:392–402.

Robins JM, Ritov Y. 1997. Toward a curse of dimensionality appropriate (CODA) asymptotic theory for semi-parametric models. Stat Med 16:285–319.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwartz J, Zanobetti A. 2000. The National Morbidity, Mortality, and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

Samet JM, Zeger SL, Kelsall JE, Xu J, Kalkstein LS. 1997. Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Air Pollutants. The Phase I.B Report of the Particle Epidemiology Evaluation Project. Special Report. Health Effects Institute, Cambridge MA.

Samoli E, Schwartz J, Analitis A, Petasakis Y, Wojtyniak B, Touloumi G, Spix C, Balducci F, Medina S, Rossi G, Sunyer J, Anderson HR, Katsouyanni K. 2003. Sensitivity analyses of regional differences in short-term effects of air pollution on daily mortality in APHEA cities. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Schwartz J. 2000. Harvesting and long term exposure effects in the relation between air pollution and mortality. Am J Epidemiol 151:440–448.

Schwartz J. 2003a. Airborne particles and daily deaths in 10 US cities. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Schwartz J. 2003b. Daily deaths associated with air pollution in six us cities and short-term mortality displacement in Boston. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Schwartz J, Zanobetti A, Bateson T. 2003. Morbidity and mortality among elderly residents of cities with daily PM measurements. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Sheppard L. 2003. Ambient air pollution and nonelderly asthma hospital admissions in Seattle, Washington, 1987–1994. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Stölzel M, Peters A, Wichmann H-E. 2003. Daily mortality and fine and ultrafine particles in Erfurt, Germany. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Wood SN. 2003. Thin plate regression splines. J Royal Stat Soc Ser B 65:95–114.

Zanobetti A, Schwartz J. 2003a. Airborne particles and hospital admissions for heart and lung disease. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Zanobetti A, Schwartz J. 2003b. Multicity assessment of mortality displacement within the APHEA2 project. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston MA.

Zeger SL, Diggle PJ. 2001. Correction: Exposure measurement error in time-series air pollution studies. Environ Health Perspect 109:A517.

Zeger SL, Thomas D, Dominici F, Samet JM, Schwartz J, Dockery DW, Cohen A. 2000. Exposure measurement error in time-series studies of air pollution: Concepts and consequences. Environ Health Perspect 108:419-426.

HEI

APPENDIX A. Highlights of Results from Short Communication Reports by First Author

ATKINSON

Project location, time	APHEA2, 8 European cities, 1990–1997
Outcomes	Hospital admissions from asthma, COPD, respiratory diseases
Exposures	Daily 24-hr average: PM_{10}, PM_{13}, black smoke, TSP; 8 hr: O_3, CO; daily 1-hr maximum: NO_2
Model adjustments	Time, temperature, humidity, holidays, day of week, trends: influenza, season[a]
	Same *df* as in original analysis

Pollutant[a], outcome	GAM-Default[b,c]	GAM Strict[b,c]	GLM[b,c]
PM_{10}			
Asthma			
0–14 years	1.2 (0.2, 2.3)	1.5 (0.1, 2.8)	See Fig 1 for city-
15–64 years	1.1 (0.3, 1.8)	1.0 (0.3, 1.8)	specific results
COPD and Asthma			
≥ 65 years	1.0 (0.4, 1.5)	1.0 (0.6, 1.4)	
Respiratory Disease			
≥ 65 years	0.9 (0.6, 1.3)	1.0 (0.7, 1.3)	
Black Smoke			
Asthma			
0–14 years	1.3 (0.3, 2.4)	1.6 (0.5, 2.6)	
15–64 years	0.7 (−0.3, 1.8)	0.7 (−0.3, 1.8)	
COPD and Asthma			
≥ 65 years	0.2 (−0.7, 1.1)	0.4 (−0.3, 1.0)	
Respiratory Disease			
≥ 65 years	0.1 (−0.7, 0.9)	0.1 (−0.7, 0.9)	

[a] Results presented for PM_{10} and black smoke only.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

BURNETT

Project location, time	Canada, 8 cities, 1986–1996; 4018 days in series
Outcomes	Nonaccidental mortality, 19.18 deaths/day
Exposures	Measurement every 6th day: $PM_{2.5}$, $PM_{10-2.5}$, PM_{10}; PM data: ~670 days; missing values imputed for some analyses
Model adjustments	LOESS for time 90-day span; temp, humidity, and change in barometric pressure 0.5 span; common weather model for all cities
	GLM: 6 knots/yr selected for 4018 days, temp and humidity 2 *df*; common and city-specific models vary. Models with ~670 days of pollutant had fewer knots. Also modeled time: 1–12 knots/yr
	GAM and GLM: day of week; Bartlett test for white noise residuals among 8 cities

Pollutant	GAM-Default[a,b]	GAM-Strict[a,b]	GLM[a,b,f]
$PM_{2.5}$[b]	1.62 (3.83)	1.44 (3.14)	0.85
$PM_{10-2.5}$[b]	1.00 (2.45)	0.83 (2.04)	0.73
PM_{10}[b]	0.87 (3.92)	0.70 (3.15)	0.53
$PM_{2.5}$[c,e]			1.10 (2.89)
$PM_{10-2.5}$[c,e]			1.16 (2.52)
PM_{10}[c,e]			0.80 (3.10)
$PM_{2.5}$[d,e]			1.05 (2.78)
$PM_{10-2.5}$[d,e]			1.21 (2.65)
PM_{10}[d,e]			0.80 (3.11)

[a] Percent change 10 µg/m^3 PM (t = ratio of estimate to SE).

[b] Lag 1.

[c] Common model (same *df* for time for all cities, 1 knot/6 months); ~670 PM days only.

[d] City-specific models (*df* for time vary by city), ~670 PM days only.

[e] Fewer *df* were required to produce white noise residuals when using only those days in which PM data were available.

[f] Lag 1, temperature, and barometric pressure 2 *df*.

Tests for heterogeneity were positive with LOESS models and negative with ns models.

Estimates varying number of knots for time in GLM: $PM_{2.5}$ 1 knot/yr 1.17% (t = 3.12); knot/2 mo 0.86% (t = 2.07); knot/mo 0.75% (t = 1.72); $PM_{10-2.5}$ GLM 1 knot/yr 1.53% (t = 3.42); knot/2 mo 0.73% (t = 1.46); knot/mo 0.49% (t = 0.91).

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

DOMINICI

20 CITIES

Project location, time	20 US cities, 1987–1994
Outcomes	Mortality: total nonaccidental, cardiovascular and respiratory (CVDRESP), and OTHER causes
Exposures	PM_{10} every 6th day measurements in most cities, 10% trimmed mean after correction for yearly averages for each monitor
Model adjustments	GAM and GLM: time 7 *df*/yr, age-specific and seasonal variation in mortality 2 *df*/yr, temp 6 *df*, dewpoint 3 *df*, influenza, indicator variables for day of week and age group

Pollutant, Outcome	**GLM**
PM_{10}, mean 0–1 lag, current day and previous day concentration-response[b]	
CVDRESP mortality	
OTHER mortality	Figs 1 and 3

88 CITIES

Project location, time	88 US cities, 1987–1994
Outcomes	Total nonaccidental mortality
Exposures	PM_{10} every 6th day measurements in most cities but nearly daily in 5 cities, 10% trimmed mean after correction for yearly averages for each monitor
Model adjustments	GAM and GLM: time 7 *df*/yr, age-specific and seasonal variation in mortality 2 *df*/yr, temp 6 *df*, dewpoint 3 *df*, influenza, indicator variables for day of week and age group

Pollutant, Outcome	**GLM[a]**
PM_{10} exposure-response, regional and overall effects	
Total mortality	See Fig 2

PHILADELPHIA

Project location, time	Philadelphia, 1974–1988
Outcomes	Deaths
Exposures	TSP
Model adjustments	Time 90 *df* (~7 *df*/yr); temp and dewpoint 6 *df*; long-term trends, medical practices, demographic characteristics, influenza

Pollutant, Outcome	**GLM**
TSP	
Total mortality	Fig 4, compared frequency domain (continuous curve) to time-scale estimates (points connected by line segments). Estimates higher with longer than with shorter time scales

[a] Hierarchical spline model: boundary knots for PM_{10} at 0 and 100 μg/m³, assuming number and location of knots the same across cities within a region. Implemented reversible jump Markov Chain Monte Carlo algorithm (RJMCMC) for model fitting. Compared three models.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

FAIRLEY

Project location, time	Santa Clara County, 1989–1996
Outcomes	Nonaccidental mortality: Total 58,440 deaths
	Cardiovascular mortality: Total 25,395 deaths
	Respiratory mortality: Total 7180 deaths
Exposure	1 in 6 days PM data: PM_{10}, $PM_{10-2.5}$, $PM_{2.5}$, $PM_{2.5}$ aug = 835 $PM_{2.5}$ observations with imputation from PM_{10} and COH; COH, PM, nitrate, sulfate, O_3 8-hr max, O_3 hrs > 60 ppb
Model adjustments	GAM: time 7 *df* (~9 df/yr) + day of year (season) 12 *df* (~2.4 *df*/yr), min temp 3 *df*, max temp 2 *df*
	GLM: time 7 *df* (~9 df/yr) + day of year (season) 4 *df* (1.4 *df*/yr), min temp 3 *df*, max temp 3 *df*; day of week

Outcome, Pollutant	GAM-Default[a,b]	GAM-Strict[a]	GLM[a]
Total mortality(lag 0)			
PM_{10}	0.08	0.078 (0.028, 0.131)	0.083 (0.029, 0.139)
$PM_{2.5}$	0.093	0.092 (0.018, 0.172)	0.080 (0.016, 0.148)
$PM_{2.5}$ aug	NA	0.083 (0.030, 0.138)	0.094 (0.034, 0.158)
$PM_{10-2.5}$	0.024	0.023 (−0.040, 0.091)	0.017 (−0.028, 0.064)
COH	0.026	0.022 (−0.010, 0.054)	0.024 (−0.012, 0.061)
NO_3	0.074	0.074 (0.025, 0.124)	0.070 (0.024, 0.117)
SO_4	0.053	0.053 (0.007, 0.101)	0.052 (0.007, 0.100)
CVD mortality (lag 0)			
PM_{10}	0.086	0.085 (0.0006, 0.170)	0.089 (0.013, 0.170)
$PM_{2.5}$	0.073	0.072 (−0.046, 0.205)	0.076 (−0.028, 0.191)
$PM_{2.5}$ aug	NA	0.103 (0.018, 0.196)	0.104 (0.025, 0.190)
$PM_{10-2.5}$	0.026	0.026 (−0.072, 0.134)	
COH	0.03	0.027 (−0.015, 0.072)	
NO_3	0.093	0.092 (0.017, 0.173)	
SO_4	0.04	0.040 (−0.028, 0.113)	
Respiratory mortality (lag 0)			
PM_{10}	0.108	0.107 (−0.037, 0.272)	0.108 (−0.034, 0.272)
$PM_{2.5}$	0.133	0.133 (−0.110, 0.442)	0.154 (−0.041, 0.389)
$PM_{2.5}$ aug		0.102 (−0.033, 0.255)	0.108 (−0.038, 0.277)
$PM_{10-2.5}$	0.156	0.156 (−0.065, 0.428)	
COH	0.1	0.097 (0.020, 0.180)	
NO_3	0.1	0.100 (−0.048, 0.272)	

[a] RR per variable increase in pollutant or other threshold level.

[b] Convergence criterion 10^{-4} instead of default of 10^{-3} in original data analysis.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

GOLD

Project location, time	Boston, May–July 1997
Outcome	Heart rate variability: heart rate and RR intervals from ECG tracings in 21 elderly subjects
Exposures	Continuous, 24-hr, and 4-hr averages. PM_{10}, $PM_{2.5}$, NO_2, SO_2, 1-hr O_3
Model adjustments	Prescription drugs, temperature, natural spline with 3 *df*

Outcome, Pollutant	GAM-Default[a]	GLM[a]
Overall heart rate		
$PM_{2.5}$ (24-hr)	−1.9 (SE 0.7)	−1.9 (SE 0.7)
r-MSSD, slow breathing[b]		
$PM_{2.5}$ (4-hr)	−6.1 (SE 2.2)	−6.0 (SE 2.2)

[a] Percent change for interquartile range: for 24-hr $PM_{2.5}$, 12 $\mu g/m^3$; for 4-hr $PM_{2.5}$, 14.35 $\mu g/m^3$.

[b] r-MSSD = square root of mean of the squared differences between adjacent normal RR intervals (intervals between adjacent R waves on an ECG tracing).

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

GOLDBERG

Project location, time	Montreal, 1984–1993
Outcomes	Nonaccidental mortality, mortality among persons with underlying disease: cancer, acute lower respiratory disease, chronic coronary artery disease, congestive heart failure, any coronary artery disease, any CVD
Exposures	every 6th day measurements data: predicted $PM_{2.5}$ and sulfate from predicted $PM_{2.5}$; COH, hourly; gaseous pollutants, continuous; Sutton sulfate, daily
Model adjustments	GAM time 73.5 *df* (~7 *df*/yr), temp and barometric pressure 9.2 *df*; GLM 88–27 *df* (~8–3 *df*/yr), temp 0–7 *df*, humidity 0–7 *df*, barometric pressure 0–6 *df*
	GAM and GLM: year; gaseous pollutants; Bartlett test for white noise residuals

Outcome, Pollutant	GAM-Default[a,b]	GAM-Strict[a,b]	GLM[a,c]
Nonaccidental deaths, < 65 years old, 3-day mean lag			
COH, ns 79 *df*	0.30	−0.30	−0.84
Predicted $PM_{2.5}$, ns 64 *df*	1.03	0.28	0.08
Nonaccidental deaths, ≥ 65 years old, 3-day mean lag			
COH, ns 79 *df*	2.57	1.92	0.85
Predicted $PM_{2.5}$, ns 64 *df*	2.68	1.97	0.70
Underlying chronic coronary disease, all ages, 3-day mean			
COH, ns 49 *df*	2.62 (0.53, 4.75)	1.94 (−0.14, 4.07)	1.10 (−1.65, 3.93)
Predicted $PM_{2.5}$, ns 36 *df*	2.20 (0.14, 4.31)	1.36 (−0.70, 3.45)	0.48 (−2.30, 3.34)
Any underlying coronary disease, all ages, 3-day mean			
COH, ns 27 df	2.99 (1.13, 4.88)	2.26 (0.41, 4.15)	1.39 (−0.95, 3.79)
Predicted $PM_{2.5}$, ns 30 *df*	1.85 (0.03, 3.70)	0.96 (−0.85, 2.80)	0.44 (−1.97, 2.91)
Underlying congestive heart failure, all ages, 3-day mean			
COH, ns 39 *df*	4.99 (2.44, 7.60)	4.14 (1.61, 6.73)	2.15 (−1.08, 5.47)
Predicted $PM_{2.5}$, ns 37 *df*	4.02 (1.61, 6.48)	2.98 (0.60, 5.42)	1.60 (−1.62, 4.92)
Underlying chronic coronary disease ≥ 65 years old			
COH (lag 0), ns 49 *df*			
temp < 25°C			6.49 (2.10, 11.06)
temp < 25°C, RH < 88			6.55 (2.15, 11.14)
Any underlying coronary disease ≥ 65 years old			
COH (lag 0), ns 27 *df*			
temp < 25°C			4.93 (1.18, 8.82)
temp < 25°C, RH < 88			5.11 (1.34, 9.01)
Underlying congestive heart failure ≥ 65 years old			
COH (lag 0), ns 39 *df*			
temp < 25°C			7.63 (2.73, 12.77)
temp < 25°C; RH < 88%			7.72 (2.80, 12.88)

[a] Percent change in daily mortality per interquartile range increase: COH, 1.85 (0.1 COH units per 327.8 linear meters); predicted $PM_{2.5}$, 9.50 µg/m³.

[b] Time 73.5 *df*.

[c] For each outcome, *df* in ns specified next to pollutant.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

HOEK

Project location, time	The Netherlands, 1986–1994
Outcomes	Total nonaccidental mortality and cause-specific mortality
Exposures	Daily measurements of black smoke, O_3, SO_2, NO_2, CO; 1992–1994 PM_{10}, $PM_{2.5}$, including sulfate and nitrate
Model adjustments	Temperature, humidity, seasons, day of week, holidays, influenza
	Time: total mortality 94 *df* (10.4 *df*/yr), cardiovascular and respiratory 60 *df* (6.67 *df*/yr)

Outcome, Pollutant	GAM-Default[a]	GAM-Strict[a]	GLM[a]
Total nonaccidental mortality			
$\quad PM_{10}$ lag 1	1.018 (1.003, 1.034)	1.019 (1.003, 1.034)	1.018 (1.002, 1.035)
$\quad PM_{10}$ average 0–6 days	1.023 (1.004, 1.041)	1.023 (1.005, 1.041)	1.019 (0.998, 1.040)
$\quad$ Black smoke lag 1	1.020 (1.010, 1.030)	1.019 (1.010, 1.029)	1.019 (1.009, 1.030)
CVD			
$\quad PM_{10}$ average 0–6 days	1.015 (0.987, 1.043)	1.015 (0.988, 1.044)	1.025 (0.995, 1.057)
$\quad$ Black smoke average 0–6 days	1.032 (1.016, 1.048)	1.031 (1.015, 1.047)	1.029 (1.010, 1.048)
$\quad$ Sulfate lag1	1.021 (0.981, 1.063)	1.022 (0.982, 1.064)	1.027 (0.984, 1.072)
$\quad$ Nitrate lag1	1.024 (0.983, 1.066)	1.025 (0.984, 1.068)	1.035 (0.990, 1.083)
COPD			
$\quad PM_{10}$ average 0–6 days	1.096 (1.014, 1.185)	1.099 (1.017, 1.188)	1.097 (1.007, 1.195)
$\quad$ Black smoke average 0–6 days	1.072 (1.026, 1.120)	1.070 (1.024, 1.117)	1.072 (1.018, 1.129)
$\quad$ Sulfate lag 1	1.115 (0.996, 1.248)	1.119 (1.000, 1.252)	1.118 (0.992, 1.261)
$\quad$ Nitrate lag 1	1.077 (0.958, 1.211)	1.081 (0.961, 1.215)	1.098 (0.966, 1.247)
Pneumonia			
$\quad PM_{10}$ average 0–6 days	1.167 (1.058, 1.287)	1.169 (1.060, 1.289)	1.176 (1.057, 1.309)
$\quad$ Black smoke average 0–6 days	1.126 (1.064, 1.192)	1.122 (1.060, 1.188)	1.137 (1.063, 1.215)
$\quad$ Sulfate lag 1	1.098 (0.954, 1.264)	1.106 (0.961, 1.273)	1.104 (0.948, 1.284)
$\quad$ Nitrate lag 1	1.202 (1.040, 1.389)	1.208 (1.045, 1.396)	1.251 (1.068, 1.466)

[a] RR (95% CI) per ~1- to 99-percentile pollutant concentration difference.

Results for mortality from myocardial infarction, arrhythmias, congestive heart failure, and cerebrovascular and thrombotic diseases also presented.

Results from two-pollutant models and seasonal analyses presented for total mortality.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

ITO

Project location, time	Detroit, 1985–1990, 1992–1994
Outcomes	Total mortality: 1985–1990, 49 deaths/day; 1992–1994, 53 deaths/day
	Circulatory mortality: 1985–1990, 25 deaths/day; 1992–1994, 25 deaths/day
	Respiratory mortality: 1985–1990, 3.7 deaths/day; 1992–1994, 4 deaths/day
	Hospital admissions among residents $\geq$ 65 years of age, 1992–1994 only: pneumonia 12/day, COPD 8/day, ischemic heart disease 22/day, dysrhythmias 7/day, heart failure 17/day, stroke 13/day
Exposures[a]	PM_{10}, $PM_{2.5}$, $PM_{10-2.5}$, TSP, TSP-PM_{10}, sulfate from TSP filters, O_3, SO_2, NO_2, CO
	1985–1990: PM_{10}, 1565 days; other pollutants: 1569–2184 days
	1992–1994: PM_{10}, $PM_{2.5}$, $PM_{10-2.5}$, 490 days; other pollutants: 344–1096 days
Model adjustments	Time 12 *df*/yr except respiratory mortality, 1985–1990, 4 *df*/yr
	Temperature LOESS ~2.5 *df*: warm (lag 0), cold (average lags 1–3), hot-humid indicator variable (> 80°F and 70%); dewpoint and barometric pressure LOESS ~2.5 *df*
	Season, day of week, influenza
	GLM: temperature and humidity 2 *df*
	Alternate GLM models: time 1–24 *df*/yr; weather variables, 3–6 *df* or varying lags

Outcome, Pollutant[a]	GAM-Default	GAM-Strict[b]	GLM[c]
1985–1990			
Total mortality			
$\quad PM_{10}$ lag 1	1.026 (1.003, 1.049)	1.104 (0.992, 1.0370)	1.010 (0.988, 1.033)
Circulatory mortality			
$\quad PM_{10}$ lag 3	1.023(0.991, 1.057)	1.010 (0.978, 1.043)	1.005 (0.974, 1.038)
Respiratory mortality			
$\quad PM_{10}$ lag 1	1.123 (1.036, 1.218)	1.114 (1.027, 1.208)	1.077 (0.994, 1.168)
1992–1994			
Total mortality			
$\quad PM_{2.5}$ lag 1	1.045 (0.991, 1.102)	1.027 (0.974, 1.083)	1.029 (0.976, 1.085)
$\quad PM_{10-2.5}$ lag 1	1.038 (0.988, 1.090)	1.031 (0.982, 1.082)	1.027 (0.979, 1.078)
$\quad PM_{10}$ lag 1	1.045 (0.989, 1.102)	1.034 (0.980, 1.091)	1.032 (0.978, 1.089)
Circulatory mortality			
$\quad PM_{2.5}$ lag 1	1.046 (0.967, 1.131)	1.032 (0.954, 1.116)	1.029 (0.952, 1.113)
$\quad PM_{10-2.5}$ lag 1	1.075 (1.000, 1.155)	1.064 (0.990, 1.143)	1.058 (0.985, 1.137)
$\quad PM10$ lag 1	1.070 (0.987, 1.160)	1.055 (0.973, 1.142)	1.050 (0.969, 1.137)
Respiratory mortality			
$\quad PM_{2.5}$ lag 0	1.033 (0.855, 1.248)	1.033 (0.854, 1.249)	1.045 (0.864, 1.265)
$\quad PM_{10-2.5}$ lag 2	1.071 (0.913, 1.257)	1.067 (0.909, 1.253)	1.061 (0.904, 1.245)
$\quad PM_{10}$ lag 0	1.080 (0.896, 1.301)	1.077 (0.893, 1.298)	1.081 (0.896, 1.303)
$\quad SO_4$ lag 3	1.066 (0.908 ,1.251)	1.056 (0.899, 1.240)	1.047 (0.892, 1.230)

Table continues next page

[a] 1220 RR estimates obtained for all pollutant models.

[b] Compared with GAM-default median drop 0.0071 per 5th to 95th percentile increase in pollutant for all outcomes combined.

[c] Compared with GAM-strict median drop 0.0028 per 5th to 95th percentile increase in pollutant for all outcomes combined.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

ITO 1992–1994 (*Continued*)

Outcome, Pollutant[a]	GAM-Default	GAM-Strict[b]	GLM[c]
Hospitalizations among those ≥ 65 years of age			
Pneumonia			
$PM_{2.5}$ lag 1	1.185 (1.053, 1.332)	1.154 (1.027, 1.298)	1.149 (1.022, 1.292)
$PM_{10-2.5}$ lag 1	1.114 (1.006, 1.233)	1.095 (0.990, 1.211)	1.107 (1.000, 1.226)
PM_{10} lag 1	1.219 (1.084, 1.372)	1.185 (1.054, 1.332)	1.190 (1.057, 1.338)
SO_4 lag 1	1.156 (1.050, 1.273)	1.128 (1.025, 1.242)	1.123 (1.025, 1.235)
COPD			
$PM_{2.5}$ lag 3	1.080 (0.933, 1.251)	1.043 (0.902, 1.207)	1.004 (0.869, 1.161)
$PM_{10-2.5}$ lag 3	1.089 (0.960, 1.236)	1.083 (0.954, 1.229)	1.103 (0.970, 1.253)
PM_{10} lag 3	1.098 (0.946, 1.274)	1.066 (0.920, 1.235)	1.047 (0.904, 1.213)
Ischemic heart disease			
$PM_{2.5}$ lag 3	1.063 (0.980, 1.153)	1.053 (0.971, 1.143)	1.043 (0.961, 1.131)
$PM_{10-2.5}$ lag 3	1.101 (1.026, 1.181)	1.098 (1.023, 1.178)	1.078 (1.004, 1.157)
PM_{10} lag 3	1.091 (1.005, 1.184)	1.082 (0.997, 1.175)	1.063 (0.980, 1.154)
Heart failure			
$PM_{2.5}$ lag 1	1.133 (1.034, 1.241)	1.117 (1.020, 1.224)	1.100 (1.004, 1.205)
$PM_{10-2.5}$ lag 0	1.050 (0.968, 1.138)	1.042 (0.962, 1.130)	1.047 (0.966, 1.135)
PM_{10} lag 0	1.099 (1.002, 1.206)	1.094 (0.997, 1.200)	1.086 (0.990, 1.191)
Stroke			
$PM_{2.5}$ lag 0	1.026 (0.925, 1.139)	1.028 (0.926, 1.140)	1.014 (0.914, 1.125)
$PM_{10-2.5}$ lag 1	1.047 (0.955, 1.148)	1.048 (0.956, 1.149)	1.054 (0.961, 1.155)
PM_{10} lag 1	1.049 (0.944, 1.165)	1.051 (0.946, 1.167)	1.045 (0.941, 1.161)

[a] 1220 RR estimates obtained for all pollutant models.

[b] Compared with GAM-default median drop 0.0071 per 5th to 95th percentile increase in pollutant for all outcomes combined.

[c] Compared with GAM-strict median drop 0.0028 per 5th to 95th percentile increase in pollutant for all outcomes combined.

KATSOUYANNI

Project location, time	29 European cities, 1990–1995 or 1993–1997
Outcomes	Total mortality
Exposures	Daily measurements of PM_{10} and black smoke
Model adjustments	Time: 10–24 *df* (~1.6–4.4 *df*/yr), temperature 2–9 *df*, season, humidity, influenza, day of week, holidays, unusual events. City-specific models

Pollutant	GAM-Default[a]	GAM-Strict[a,b]	GLM[a,b]
Average, 0–1 lag			
PM_{10}[c]	0.000617 (0.000106)	0.000593 (0.000103)	0.000410 (0.000091)
BS	0.000576 (0.000133)	0.000552 (0.000131)	0.000363 (0.000112)

[a] β (SE). City-specific random effects models. Percent increase in mortality per 10 μg/m^3.

[b] Percent difference in random mean estimate compared with GAM-default. For PM_{10}, 21 cities: GAM-strict −3.9, GLM −33.6, penalized splines −10.7. For black smoke, 14 cities: GAM-strict −4.2, GLM −37.1.

[c] Penalized splines β 0.000550 (SE 0.000097).

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

KLEMM

Project location, time	6 cities, US 1979–1988
Outcomes	Mortality from nonaccidental causes, COPD, ischemic heart disease, pneumonia
Exposures	Daily or every other day measurements: $PM_{2.5}$, PM_{10}, PM_{15}, coarse mass fraction, SO_4
Model adjustments	For $PM_{2.5}$: GAM-default time 36.8 *df* (~3.7 *df*/yr), dewpoint 3.5 *df*; GAM-strict: time 37 *df*, day of week 6 *df*, weather 7 *df*; GLM: time 4–12 knots/yr; additional GAMs with LOESS: time > 160 *df*; GLM-natural splines: time > 250 *df*.
	Degrees of freedom for time varied slightly for other particulate mass concentrations

Outcome, Pollutant	GAM-Default (time 36.8 *df*)	GAM-Strict (time 37 *df*)	GLM (time 38 *df*)	GLM (12 knots/yr)
Nonaccidental mortality, 2-day mean lag				
$PM_{2.5}$	1.3 (0.9, 1.7)	1.2 (0.8, 1.6)	0.8 (0.4, 1.3)	0.4 (−0.1, 0.8)
Coarse mass	0.4 (−0.2, 0.9)	0.3 (−0.2, 0.9)	0.1 (−0.5, 0.7)	−0.2 (−0.8, 0.5)
PM_{15} or PM_{10}	0.8 (0.5, 1.0)	0.7 (0.4, 1.0)	0.4 (0.1, 0.8)	0.1 (−0.2, 0.5)
SO_4	1.6 (0.9, 2.4)	1.5 (0.8, 2.3)	1.0 (0.1, 1.9)	0.4 (−0.5, 1.3)
Ischemic heart disease, 2-day mean lag				
$PM_{2.5}$	2.0 (1.3, 2.7) $t = 5.50$	1.8 (1.1, 2.5) $t = 5.18$	1.5 (0.7, 2.3) $t = 3.75$	0.8 (0.0, 1.6) $t = 1.96$

LE TERTRE

Project location, time	8 European cities, ~1990– ~1997, 930–2829 days
Outcomes	Hospital admissions for cardiac diagnoses and stroke
Exposures	Daily measurements PM_{10} (PM_{13} in Paris)
Model adjustments	Time, season, temperature, humidity, influenza, day of week, holidays, unusual events
	Same *df* as in original analysis

Outcome	GAM-Default[a,b]	GAM-Strict[a,b]	GLM[a,b]
Cardiac (≥ 65 years)	0.7 (0.4, 1.0)	0.7 (0.4, 0.9)	See figures for city-specific results
Ischemic heart disease	0.8 (0.3, 1.2)	0.7 (0.3, 1.2)	
Stroke	0.0 (−0.3, 0.3)	0.0 (−0. 3, 0.3)	
Cardiac[c]	0.0006333 (0.0001342)	0.0006685 (0.0001374)	0.00091537 (0.0001565)

[a] Random effects models.

[b] RR (95% CI) per 10 μg/m^3 in PM_{10} or PM_{13}.

[c] β (SE); penalized splines β 0.00068102 (SE 0.00015489).

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

---

MAR

Project location, time	Maricopa County AZ, 1995–1997
Outcomes	CVD mortality among residents 65–100 years of age
Exposures	Daily measurements PM_{10}, $PM_{2.5}$, PM coarse fraction, OC, EC, TC, CO, vehicle exhaust, road dust, soil, vegetative burning, local source of SO_2, regional SO_4
Model adjustments	GAM: time 10 *df* (~3.3 *df*/yr); temperature lagged 1 day, 2 *df*; humidity lag 0, 2 *df* GLM: time 10 *df* (~3.3 *df*/yr); temperature lagged 1 day, 2 *df*; humidity lag 0, 3 *df*

Pollutant	GAM-Default[a]	GAM-Strict[a]	GLM
PM_{10}			
lag 0	1.05 (1.01, 1.09)	1.05 (1.01, 1.09)	1.05 (1.01, 1.09)
lag 1	1.04 (1.00, 1.08)	1.04 (1.00, 1.08)	1.04 (1.00, 1.08)
lag 3	1.03 (0.99, 1.07)	1.03 (0.99, 1.07)	1.03 (0.99, 1.07)
lag 4	1.03 (0.99, 1.07)	1.03 (0.99, 1.07)	1.03 (0.99, 1.08)
$PM_{2.5}$			
lag 0	1.03 (0.99, 1.08)	1.03 (0.99, 1.07)	1.03 (0.99, 1.08)
lag 1	1.06 (1.02, 1.10)	1.06 (1.02, 1.10)	1.06 (1.01, 1.11)
lag 3	1.04 (1.00, 1.08)	1.04 (1.00, 1.08)	1.04 (0.99, 1.09)
lag 4	1.05 (1.01, 1.09)	1.05 (1.01, 1.09)	1.05 (1.01, 1.09)
Coarse fraction ($PM_{10-2.5}$)			
lag 0	1.05 (1.01, 1.09)	1.05 (1.01, 1.08)	1.05 (1.01, 1.09)
lag 1	1.03 (0.99, 1.07)	1.03 (0.99, 1.07)	1.03 (0.99, 1.07)
lag 3	1.02 (0.98, 1.05)	1.02 (0.98, 1.05)	1.02 (0.98, 1.06)
lag 4	1.02 (0.98, 1.06)	1.02 (0.98, 1.06)	1.02 (0.99, 1.06)
SO_4[b]			
lag 0	1.06 (1.00, 1.12)	1.06 (1.00, 1.12)	1.06 (1.00, 1.13)
lag 1	1.02 (0.96, 1.07)	1.02 (0.96, 1.08)	1.02 (0.96, 1.08)
lag 3	1.00 (0.95, 1.06)	1.00 (0.95, 1.06)	1.01 (0.95, 1.07)
lag 4	1.00 (0.95, 1.05)	1.00 (0.95, 1.05)	1.01 (0.95, 1.07)
Motor vehicle exhaust and resuspended road dust			
lag 0	1.01 (0.97, 1.06)	1.01 (0.97, 1.05)	1.01 (0.96, 1.06)
lag 1	1.06 (1.01, 1.10)	1.06 (1.01, 1.10)	1.05 (1.00, 1.10)
lag 3	1.01 (0.97, 1.06)	1.01 (0.97, 1.06)	1.00 (0.95, 1.05)
lag 4	1.02 (0.98, 1.07)	1.02 (0.98, 1.07)	1.01 (0.96, 1.06)

---

[a] RR estimates per interquartile range increase: PM_{10} 24.88 µg/m³, $PM_{2.5}$ 8.47–8.52 µg/m³, PM coarse fraction 18.39 µg/m³, regional sulfates 1.38 units.

[b] Regional sulfates.

Results for lags 0 to 4 presented for all exposures.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

MOOLGAVKAR

Project location, time	Los Angeles County CA and Cook County IL, 1987–1995
Outcomes	Nonaccidental mortality, mortality from CVD and COPD
	Hospital admissions for CVD and COPD
Exposures	PM every 6th day measurements, PM_{10}, CO, SO_2, NO_2, $PM_{2.5}$, PM_{10}, 2638 and 2788 days in 2 counties
Model adjustments	Time 30 *df* (~3.7 *df*/yr) and 100 *df* (~12.5 *df*/yr), temperature 6 *df*
	For total mortality:
	Los Angeles County temperature lagged 1 day, relative humidity, same day, 6 *df*;
	Cook County temperature lagged 2 days, relative humidity lagged 3 days, 6 *df*

Outcome, pollutant	GAM-Default[a,b]	GAM-Strict[a]	GLM[a]
Total mortality, LA County			
PM_{10} 30 *df*	0.48 (2.55)	0.47 (2.51)	0.45 (2.07)
PM_{10} 100 *df*	0.38 (2.27)	0.36 (2.08)	0.34 (1.60)
$PM_{2.5}$ 30 *df*	0.60 (1.99)	0.59 (1.96)	0.55 (1.50)
$PM_{2.5}$ 100 *df*	0.13 (0.48)	0.10 (0.35)	−0.01 (−0.02)
CVD mortality, LA County			
PM_{10} 30 *df*	0.89 (3.10)	0.88 (3.08)	Not reported
PM_{10} 100 *df*	0.80 (2.97)	0.77 (2.87)	0.77 (2.30)
$PM_{2.5}$ 30 *df*	1.03 (2.30)	1.03 (2.30)	Not reported
$PM_{2.5}$ 100 *df*	0.79 (1.85)	0.80 (1.88)	0.69 (1.34)
COPD mortality, LA County			
PM_{10} 30 *df*		0.60 (0.75)	Not reported
PM_{10} 100 *df*		0.30 (0.38)	0.24 (0.24)
$PM_{2.5}$ 30 *df*		0.38 (0.30)	Not reported
$PM_{2.5}$ 100 *df*		0.86 (0.69)	0.20 (0.13)
Total mortality, Cook County			
PM_{10} 30 *df*		0.43 (3.86)	Not reported
PM_{10} 100 *df*		0.38 (3.65)	0.37 (3.28)
CVD mortality, Cook County			
PM_{10} 30 *df*		0.24 (1.28)	Not reported
PM_{10} 100 *df*		0.21 (1.22)	0.21 (1.14)
COPD mortality, Cook County			
PM_{10} 30 *df*		0.95 (1.89)	Not reported
PM_{10} 100 *df*		0.88 (1.76)	0.94 (1.75)
CVD hospitalizations, LA County, ≥ 65 years of age			
PM_{10} 30 *df*		0.33 (1.37)	Not reported
PM_{10} 100 *df*		~0 (−0.06)	−0.07 (−0.25)
$PM_{2.5}$ 30 *df*		1.39 (3.88)	Not reported
$PM_{2.5}$ 100 *df*		1.13 (3.32)	1.20 (2.88)
CVD hospitalizations, Cook County, ≥ 65 years of age			
PM_{10} 100 *df*	0.51 (4.07)	0.49 (3.95)	0.51 (3.77)
COPD hospitalizations, LA County, ≥ 65 years of age			
PM_{10} 30 *df*		0.26 (2.00)	Not reported
PM_{10} 100 *df*		0.14 (1.06)	0.13 (0.83)
$PM_{2.5}$ 30 *df*		1.19 (2.41)	Not reported
$PM_{2.5}$ 100 *df*		0.75 (1.71)	0.77 (1.45)

[a] Percent change per 10 µg/m³ PM_{10}; *t*-statistic in parentheses; lag 1 for Cook County and Los Angeles County, $PM_{2.5}$; lag 2 for Los Angeles County, PM_{10}.

[b] PM_{10} results published in 2000 were reported per 25 µg/m³ PM concentration.

Mortality and hospital admissions also reported for other pollutants.

Results from two-pollutant models also presented.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

OSTRO

Project location, time	Coachella Valley CA, 1989–1998
Outcomes	CVD mortality 0–10 deaths/day, mean 2.7 deaths/day, 3677 days
Exposures	PM every 6th day measurements, PM_{10} 3011 days; $PM_{10-2.5}$ 789 days + 2990 predicted; $PM_{2.5}$, 1041 days, 1996–1998; O_3; CO
Model adjustments	Examined 10–60 *df* for time, temperature, humidity and dewpoint up to 4 day lags, day of week

Outcome, pollutant	GAM-Default[a]	GAM-Strict[a]	GLM[a]
CVD mortality			
Lag 0			
PM_{10}	1.09 (SE 0.38) [*t* 2.83]	1.07 (SE 0.38) [*t* 2.82]	0.99 (SE 0.38) [*t* 2.57]
$PM_{10-2.5}$	1.23 (SE 0.43) [*t* 2.88]	1.16 (SE 0.45) [*t* 2.58]	1.08 (SE 0.46) [*t* 2.36]
$PM_{2.5}$	−5.60 (SE 3.38) [*t* −1.66]	−5.80 (SE 3.40) [*t* −1.71]	−5.74 (SE 3.45) [*t* −1.66]
Lag 1			
PM_{10}	0.75 (SE 0.38) [*t* 1.97]	0.74 (SE 0.38) [*t* 1.95]	0.69 (SE 0.39) [*t* 1.77]
$PM_{10-2.5}$	0.86 (SE 0.43) [*t* 2.01]	0.89 (SE 0.45) [*t* 1.98]	0.82 (SE 0.45) [*t* 1.81]
$PM_{2.5}$	−1.42 (SE 3.26) [*t* −0.44]	−1.30 (SE 3.30) [*t* −0.39]	−1.57 (SE 3.36) [*t* −0.47]
Lag 2			
PM_{10}	0.73 (SE 0.38) [*t* 1.91]	0.72 (SE 0.38) [*t* 1.89]	0.67 (SE 0.39) [*t* 1.73]
$PM_{10-2.5}$	0.83 (SE 0.43) [*t* 1.94]	0.90 (SE 0.44) [*t* 2.05]	0.84 (SE 0.45) [*t* 1.88]
$PM_{2.5}$	−2.29 (SE 3.22) [*t* −0.71]	−2.30 (SE 3.20) [*t* −0.72]	−2.40 (SE 3.30) [*t* −0.74]
2-day lag, time 60 *df*, temperature 20 *df*			
$PM_{10-2.5}$	0.83 (0.43) original	0.97 (0.45)	0.93 (0.48)
Average of 0, 1, 2 day lags[b]			
$PM_{10-2.5}$		2.11 (0.65)	

[a] Coefficient (SE) per interquartile range (*t*-statistic also presented).

[b] Temperature 49–125°F (mean 88.7). Omitting 5% of days with high wind or temperature, time 10 *df*, unlagged temperature, day of week.

Results for lags 3 and 4 and from some two-pollutant models also presented.

[c] B-spline with 60 *df* for time β 0.94 (SE 0.47)

SAMOLI

Project location, time	APHEA2, 7 Western Europe and 5 Central Europe cities, 1980–1992
Outcomes	Total mortality
Exposures	Daily measurements of black smoke in 8 cities and SO_2 in 12 cities
Model adjustments	Time 12–57 *df*, temperature 2–9 *df*, humidity 5 *df* or linear, influenza, day of week, holidays, unusual events

Pollutant (sin/cos)[a]	GAM-Default[a]	GAM-Strict[a]	GLM[a]
Black smoke			
West [0.00057 (0.00008)]	0.00060 (0.00007)	0.00054 (0.00007)	0.00032 (0.00008)
East [0.00012 (0.00005)]	0.00045 (0.00006)	0.00041 (0.00006)	0.00020 (0.00007)

[a] Coefficient (SE). City-specific random effects models.

Percent difference for black smoke between GLM-natural splines and GAM-strict and GLM-natural splines and sin/cos: west −41.9 and −44.6; east −51.8 and +65.6, respectively.

Results for SO_2 also presented.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

---

SCHWARTZ 10 CITIES

Project location, time	10 US cities, 1986–1993		
Outcomes	Mortality		
Exposures	Daily measurement PM_{10}		
Model adjustments	Autocorrelation of residuals for white noise; same *df* as in original analyses (mean 0–1 lag)		

Outcome	GAM-Default[a]	GAM-Strict[a]	GLM[a]
Total mortality			
PM_{10} unconstrained distributed lag[b]	1.29 (0.13)	1.13 (0.14)	
PM_{10} summer only, mean 0–1 lag	0.67 (0.48, 0.86)	0.68 (0.49, 0.87)	0.52 (0.31, 0.73)
PM_{10} winter only, mean 0–1 lag	0.66 (0.45, 0.87)	0.61 (0.42, 0.83)	0.58 (0.35, 0.80)

[a] Percent change (SE or 95% CI) per 10 µg/m³ PM_{10}.

[b] Penalized splines 1.03 (0.14).

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

SCHWARTZ BOSTON AND SIX CITIES

Project location, time	Boston MA, 1979–1986; Six-Cities Study, ~1979– ~1988
Outcomes	Boston: mortality displacement; Six Cities: daily mortality
Exposures	Boston: daily $PM_{2.5}$ measurements; Six Cities: daily measurements PM_{10}, $PM_{2.5}$, traffic, coal, dirt
Model adjustments	Boston: seasonal and trend decomposition LOESS procedure with a window of 120 days. Windows of 15, 30, 45, and 60 days examined as midscale components
	Six Cities: GAM, time 36 *df* (~4 *df*/yr), temperature and humidity 3.6 *df*; GLM, same as GAM except for 1 *df* added to each variable; B-splines, thin-plate splines, and penalized splines modeled to approximate original models

Outcome: Boston	GAM-Default[a]	GAM-Strict[a]	GLM[a]
Ischemic heart disease			
15 day	0.00276 (0.00050)		0.00276 (0.00050)
30 day	0.00370 (0.00050)		0.00363 (0.00051)
45 day	0.00404 (0.00049)		0.00400 (0.00050)
60 day	0.00467 (0.00043)		0.00476 (0.00044)
Pneumonia			
15 day	0.00287 (0.00122)		0.00295 (0.00124)
30 day	0.00666 (0.00108)		0.00641 (0.00109)
45 day	0.01068 (0.00108)		0.01082 (0.00111)
60 day	0.01160 (0.00106)		0.01129 (0.00108)
COPD[b]			
15 day	0.00644 (0.00189)		0.00659 (0.00192)
30 day	0.00358 (0.00178)		0.00371 (0.00179)
45 day	0.00120 (0.00178)		0.00138 (0.00181)
60 day	−0.00050 (0.00166)		−0.00025 (0.00170)

Outcome: Six Cities	GAM-Default[b]	GAM-Strict[b]	GLM[b]
Mortality (0–1 mean lag)			
$PM_{2.5}$	0.146 (0.020)	0.137 (0.020)	0.129 (0.021)

[a] Coefficient (SE).

[b] $\beta \times 100$ (SE). Penalized splines 0.113 (0.022), time 50 knots, weather factors 10 knots. B-splines 0.017 (0.020). Thin-plate splines 0.104 (0.022).

City-specific penalized splines results for coarse mass and other exposures also presented.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

SHEPPARD

Project location, time	Seattle, 1987–1994
Outcomes	Asthma hospitalizations (2.7 admissions/day)
Exposures	$PM_{2.5}$ (72% and 81% missing values per monitor), PM_{10} (4%, 31%, and 40% missing values), coarse mass, O_3, SO_2, CO. Missing values imputed
Model adjustments	Time 64 *df* (9.15 *df*/yr), temperature (GLM 4 *df*), day of week

Pollutant	GAM-Default[a]	GAM-Strict[a]	GLM[a]
Lag 1			
PM_{10}	1.05 (1.02,1.08)	1.04 (1.01, 1.07)	1.03 (1.00, 1.06)
$PM_{2.5}$	1.04 (1.02, 1.07)	1.04 (1.01, 1.06)	1.03 (1.01, 1.06)
$PM_{10-2.5}$	1.04 (1.01, 1.07)	1.02 (1.00, 1.05)	1.02 (0.99, 1.04)

[a] RR (95% CI) per interquartile range: PM_{10} 19 µg/m³; $PM_{2.5}$ 11.8 µg/m³; $PM_{10-2.5}$ 9.3 µg/m³.

STÖLZEL

Project location, time	Erfurt, Germany, 1995–1998
Outcomes	Mortality
Exposures	Daily measurement PM_{10}, $PM_{2.5}$, PM ultrafine, number concentration of particles (NC), SO_2, NO_2, CO
Model adjustments	Trends (14 *df*/yr), temperature, humidity, influenza, day of week

Pollutant	GAM-Default[a]	GAM-Strict[a]	GLM[a]
$NC_{0.01-2.5}$ (lag 4)	1.041 (0.991, 1.093)	1.040 (0.991, 1.092)	1.043 (0.989, 1.101)
$MC_{0.01-2.5}$ (lag 0)	1.031 (1.000, 1.063)	1.030 (0.999, 1.062)	1.029 (0.993, 1.066)
$PM_{2.5}$ (lag 3)	0.970 (0.941, 1.000)	0.970 (0.941, 0.999)	0.975 (0.941, 1.010)
PM_{10} (lag 0)	1.035 (1.001, 1.069)	1.034 (1.001, 1.069)	1.029 (0.990, 1.069)
TSP (lag 1)	1.023 (0.981, 1.067)	1.023 (0.981, 1.067)	1.014 (0.968, 1.062)

[a] RR (95% CI) per interquartile range. See tables in Stölzel et al for interquartile ranges.

Results for several lags and pollutants presented. MC = mass concentration of particles.

APPENDIX A (*Continued*). Highlights of Results from Short Communication Reports by First Author

ZANOBETTI HOSPITAL ADMISSIONS

Project location, time	14 US cities, 1985–1994
Outcomes	Hospitalizations for CVD and respiratory diseases
Exposures	Daily PM_{10}
Model adjustments	City-specific models. GAM: time 350/2064 to 130/2542 span, temperature 0.4–0.9 span, temperature lag 1 0.4–0.9 span, humidity 0.4–0.7 span, barometric pressure 0.4–0.7 span, day of week 0.5–0.6 span. GLM: time 10.5–36.2 *df* (2.1–6.0 *df*/yr), temperature 1.5–4.7 *df*, temp lag 1 1.5–4.7 *df*, humidity 2.1–4.4 *df*, barometric pressure 2.4–5.0 *df*, day of week 2.6–6.0 *df*

Pollutant, Outcome	GAM-Default	GLM	Penalized Spline
PM_{10} on days < 50 µg/m³ only (2-day mean lag)			
CVD[a]		1.45 (1.12,1.78)	1.32 (0.77,1.87)
COPD[a]		2.60 (1.40,3.81)	2.21 (1.02,3.41)
Pneumonia[a]		2.46 (1.16,3.78)	1.06 (0.06,2.07)
CVD[b]	−50.5 (SE 14.6)	−37.65 (SE 17.10)	−38.70 (SE 16.14)

[a] Percent change (95% CI) in hospitalizations per 10 µg/m³ increase.

[b] Percent change in hospitalizations per interquartile range increase in air conditioned homes. Adjusted for winter peaks in cities.

ZANOBETTI APHEA2

Project location, time	APHEA2, 10 European cities, 1992–1996
Outcomes	Mortality displacement
Exposures	Daily PM_{10}
Model adjustments	City-specific models: time 3–23 *df* (1–4.6 *df*/yr), temperature lag 0 2–6 *df*, temperature mean 2–3 days 2–6 *df*, humidity lag 0 or mean lag 1–2 1.2–2 *df*

Outcome, pollutant	GAM-Default[a]	GAM-Strict[a]	Penalized Spline[a]
Total mortality			
PM_{10} mean lag 0–1	0.70 (0.14, *t* 5.13)	0.67 (0.14, *t* 4.80)	0.57 (0.15, *t* 3.82)
PM_{10} distributed lags[b]	1.61 (0.30, *t* 5.32)	1.45 (0.30, *t* 4.79)	1.08 (0.40, *t* 2.73)

[a] $\beta \times 1000$ (SE, *t*).

[b] 4th degree polynomial distributed lag models with 40 days.

APPENDIX B. Possibility of Inadequate Confounder Control by Minimizing AIC

In this appendix we use explicit examples to show that choosing degrees of freedom by minimizing AIC does not guarantee adequate control of confounding. To do so, we follow Dominici (*http://biosun01.biostat.jhsph.edu/~fdominic/*) in reducing the problem to its simplest form: a linear model with known variance rather than a log-linear model with overdispersed Poisson variance and unknown overdispersion. Our goal is estimation of the parameter β with X pollution and V either time or temperature in the following semiparametric regression model: For $t = 1, \dots, n,$

$$Y_t = \beta X_t + f(V_t) + e_t \tag{1}$$

and

$$X_t = g(V_t) + \xi_t, \tag{2}$$

where e_t are independent $N(0, \sigma^2), \xi_t$ are independent $N(0, \sigma_\xi^2)$, β is an unknown parameter representing the magnitude of the pollution effect, $f(\cdot)$ and $g(\cdot)$ are unknown functions, and $V_t = t$ if V is time. Given some complete set of known basis functions (say a b-spline basis) $\{h_k(\cdot); k = 1, 2, \dots\}$, we have the expansion $f(V_t) = \sum_{k=1}^{\infty} \delta_k h_k(V_t)$ where the δ_k are unknown coefficients. We suppose that we have selected basis functions with the property that if $f(\cdot)$ is quite smooth with few wiggles then δ_k is either very small or equal to 0 when k is large. In the air-pollution literature, allotting m degrees of freedom to $f(\cdot)$ means we approximate

$$\sum_{k=1}^{m} \delta_k h_k(\cdot)$$

and estimate β by the ordinary least squares (OLS) estimate $\hat{\beta}_m$ in the model

$$E[Y_t \mid X_t, V_t] = \beta X_t + \sum_{k=1}^{m} \delta_k h_k(V_t). \tag{3}$$

Selecting the number of degrees of freedom thus means selecting m.

AIC selects m as the $\hat{m}$ that minimizes $\sum (Y_t - \hat{Y}_{m,t})^2 + \hat{\sigma}^2 2m$ where $\hat{\sigma}^2$ is an estimate σ^2 and $\hat{Y}_{m,t}$ is the predicted value of Y_t under the fitted model. In the interest of simplifying even further, we consider the special case in which σ^2 is known, in which case AIC minimizes the statistic $\sum (Y_t - \hat{Y}_{m,t})^2 + \sigma^2 2m$. This statistic is identical to the Mallow Cp statistic and is also the Stein unbiased estimator of risk (ie, of mean squared error).

In the literature AIC is used in one of two different ways. In method 1, X_t is forced into the model when choosing m. In method 2, the X_t term is left out of the model when choosing m. As discussed in Robins and Greenland (1986) and demonstrated in the examples below, the method in which X_t is forced into the model when choosing $\hat{m}$ is much more likely to result in bias than the method in which X_t is not in the model when choosing $\hat{m}$. However, even this latter method can result in substantial bias under certain data-generating mechanisms. In fact, Ritov and Bickel (1990) showed that for continuous V_t there was no strictly data-based method by which one can decide whether a sufficient number of terms m have been selected to either make $\hat{\beta}_m$ nearly unbiased for β. Indeed they show for continuous V_t that there can be no strictly data-based method that is guaranteed to yield either (1) a nearly unbiased estimator of β or (2) a confidence interval for β that is both narrow enough to be substantively useful for risk assessment and yet wide enough to guarantee that it covers the true value of β at least 95% of the time. In this sense estimation of β is an ill-posed problem that can only be solved by addition of external a priori information. Specifically, Ritov and Bickel (1990) and Robins and Ritov (1997) show that accurate prior knowledge regarding the maximum wiggliness of the less wiggly of the functions $f(\cdot)$ and $g(\cdot)$ is necessary in order to know how large a value of m must be slected for $\hat{\beta}_m$ to be nearly unbiased and for a confidence interval (with substantively useful width) to cover the true value of β at least 95% of the time.

AIC

AIC Method 1

We show by use of an unrealistic extreme example that using AIC does not logically guarantee a small bias under method 1. We argue conditonally on $\{V_t; t = 1, \dots, T\}$ where T is the length of the time series, and we often take the $X_t -$ error variance σ_ξ^2 to be zero in our examples. Suppose that the $h_k(V_t)$ constitute a (conditional) orthonormal basis in the sense that the sum $\sum h_{k*}(V_t)h_k(V_t)$ equals 0 whenever $k* \neq k$ and is 1 when $k* = k$, δ_k is quite large for $k \leq 11$, $\delta_k = 0$ for $k > 11$, X_t and $h_{11}(V_t)$ are perfectly correlated (which implies that $\sigma_\xi^2 = 0$), β = 0, and δ_{11} differs sufficiently from 0 that a χ_1^2 test for adding $h_{11}(V_t)$ would have been significant at an α level of less than 10^{-9} level with probability essentially 1, had X_t not been forced into the model. In this setting, AIC selects $m = 10$ with probability more than 0.96 (the probability a χ_1^2 variable is less than 4) and m equal to 12 or more with probability about 0.04, so we are usually fitting the model

$$E[Y_t \mid X_t, V_t] = \beta X_t + \sum_{k=1}^{10} \delta_k h_k(V_t).$$

289

Note that because under method 1, X_t is always forced into the model, the perfect correlation between X_t and $h_{11}(V_t)$ is the sole reason that AIC did not choose $m \geq 11$ with probability essentially 1. Further the fraction 0.96 of the time that $m = 10$ is selected, a test of $\beta = 0$ based on the selected model

$$E[Y_t \mid X_t, V_t] = \beta X_t + \sum_{k=1}^{10} \delta_k h_k(V_t)$$

will incorrectly reject $\beta = 0$ with a P value less than 10^{-9} because the effect of $h_{11}(V_t)$ will be inappropriately attributed to X_t. Additionally, conditional on $m = 10$ being selected, $\hat{\beta}_{10}$ will have an expected value of approximately $g\delta_{11}$ rather than its true value $\beta = 0$, where g is the slope of the regression of X_t on $h_{11}(V_t)$. [The fraction 0.04 of the time the selected m exceeds 12, a test of $\beta = 0$ would not reject and no estimate of β would be reported as the design matrix would not be invertible since the model would include the perfectly correlated variables X_t and $h_{11}(V_t)$.]

In this example, had X_t not been forced in the model when selecting m, then with probability essentially 1, m would be chosen greater than 10, a test of $\beta = 0$ would not reject and no estimate of β would be reported as the design matrix would not be invertible [since the model would include the perfectly correlated variables X_t and $h_{11}(V_t)$]. This implies that we also could not reject the possibility that the pollution effect β was large. That is, had X_t not been forced in the model, we would no longer have had a biased estimate of β. Rather, we would have appropriately recognized that we would have no power to detect a pollution effect because of the perfect correlation of X_t with $h_{11}(V_t)$, at least in the absence of sharp prior knowledge that the true coefficient δ_{11} of $h_{11}(V_t)$ was either zero or at least very small.

AIC Method 2

We now use a different unrealistic extreme example to show that AIC does not logically guarantee a small bias even under method 2. Suppose we modify the previous example by assuming $h_{200}(V_t)$ is perfectly correlated with X_t and that $\delta_{200} \neq 0$, $\delta_k = 0$, and $k \notin \{1, 2, \dots, 10\} \cup 200$. We continue to assume $\beta = 0$. Suppose the χ^2_{190} statistic for adding 190 more variables (to make $m = 10 + 190$) has a value less than $2(190) = 380$ but greater than 20 with probability essentially 1 (so that AIC under method 2 is minimized at $\hat{m}$ less than 200 with probability nearly 1). Then a test of $\beta = 0$ based on the AIC selected model

$$E[Y_t \mid X_t, V_t] = \beta X_t + \sum_{k=1}^{\hat{m}} \delta_k$$

will incorrectly reject $\beta = 0$ with a P value less than that associated with a χ^2_1 variable being 20 or greater! Further $\hat{\beta}_{\hat{m}}$ will have an expected value $\sigma_{200}g$ rather than the true $\beta = 0$ where g is the slope of the regression of X_t on $h_{200}(V_t)$ on X_t.

Note that in this example, a difficulty is that AIC is an inclusive stepwise procedure in the sense that $h_k(V_t)$ cannot be selected into the model unless $h_{k*}(V_t)$ is selected for all $k* < k$. Had we used "all possible subset selection" methods this difficulty could have been somewhat ameliorated. However, when using "all possible subset selection," a penalty greater than the AIC penalty of $\sigma^2 2m$ must be used to account for the selection bias that comes from searching over all subsets. More importantly, even when using a correct "all-subset selection" method, examples can be constructed of data-generating processes under which severe confounding bias in the estimation of β can still occur.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike information criterion
APHEA	Air Pollution and Heath: A European Approach
BS	black smoke
CI	confidence interval
COH	coefficient of haze
COPD	chronic obstructive pulmonary disease
CVD	cardiovascular disease
df	degrees of freedom
EC	elemental carbon
ECG	electrocardiography
EPA	Environmental Protection Agency (US)
GAM	generalized additive model
GAM-default	GAM with default convergence criteria
GAM-strict	GAM with stricter convergence criteria
GLM	generalized linear model
GLM-natural splines	GLMs with natural cubic splines
IHD	ischemic heart disease
LOESS	locally weighted smoothers
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
ns	natural spline
OC	organic carbon
PM	particulate matter
PM_{10}	PM less than 10 µm in aerodynamic diameter

$PM_{10-2.5}$ PM 2.5–10 µm in aerodynamic diameter

$PM_{2.5}$ PM less than 2.5 µm in aerodynamic diameter

RR   relative risk

TC   total carbon

TSP   total suspended particles

RELATED HEI PUBLICATIONS

TIME-SERIES STUDIES

Report Number	Title	Principal Investigator	Publication Date*

Research Reports

98	Daily Mortality and Fine and Ultrafine Particles in Erfurt, Germany	H-E Wichmann	2000
	Part I: Role of Particle Number and Particle Mass		
97	Identifying Subgroups of the General Population That May Be Susceptible to Short-Term Increases in Particulate Air Pollution: A Time-Series Study in Montreal, Quebec	MS Goldberg	2000
95	Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations	M Lippmann	2000
94	The National Morbidity, Mortality, and Air Pollution Study	JM Samet	2000
	Part I: Methods and Methodologic Issues		
	Part II: Morbidity and Mortality from Air Pollution in the United States		

Copies of these reports can be obtained from the Health Effects Institute and many are available at www.healtheffects.org.

293

HEI

HEALTH EFFECTS INSTITUTE

Charlestown Navy Yard
120 Second Avenue
Boston MA 02129-4533 USA
+1-617-876-6700

SPECIAL REPORT

May 2003